Past and Future of Digital Watermarking

Santa Clara IEEE SP Chapter

Ton Kalker

Hewlett-Packard Labs

Executive Summary

- Digital watermarking is a well-understood technique for hidden communication.
- However, despite initial high hopes, digital watermarking has not had a major impact on copy right, copy protection and Digital Rights Management.
- One of the main reasons for this state of affair is the poorly understood security aspects of digital watermarking. Improving this understanding is a major open problem.
- Any claim that watermarking is relevant for copyright protection should be taken with a grain of salt.

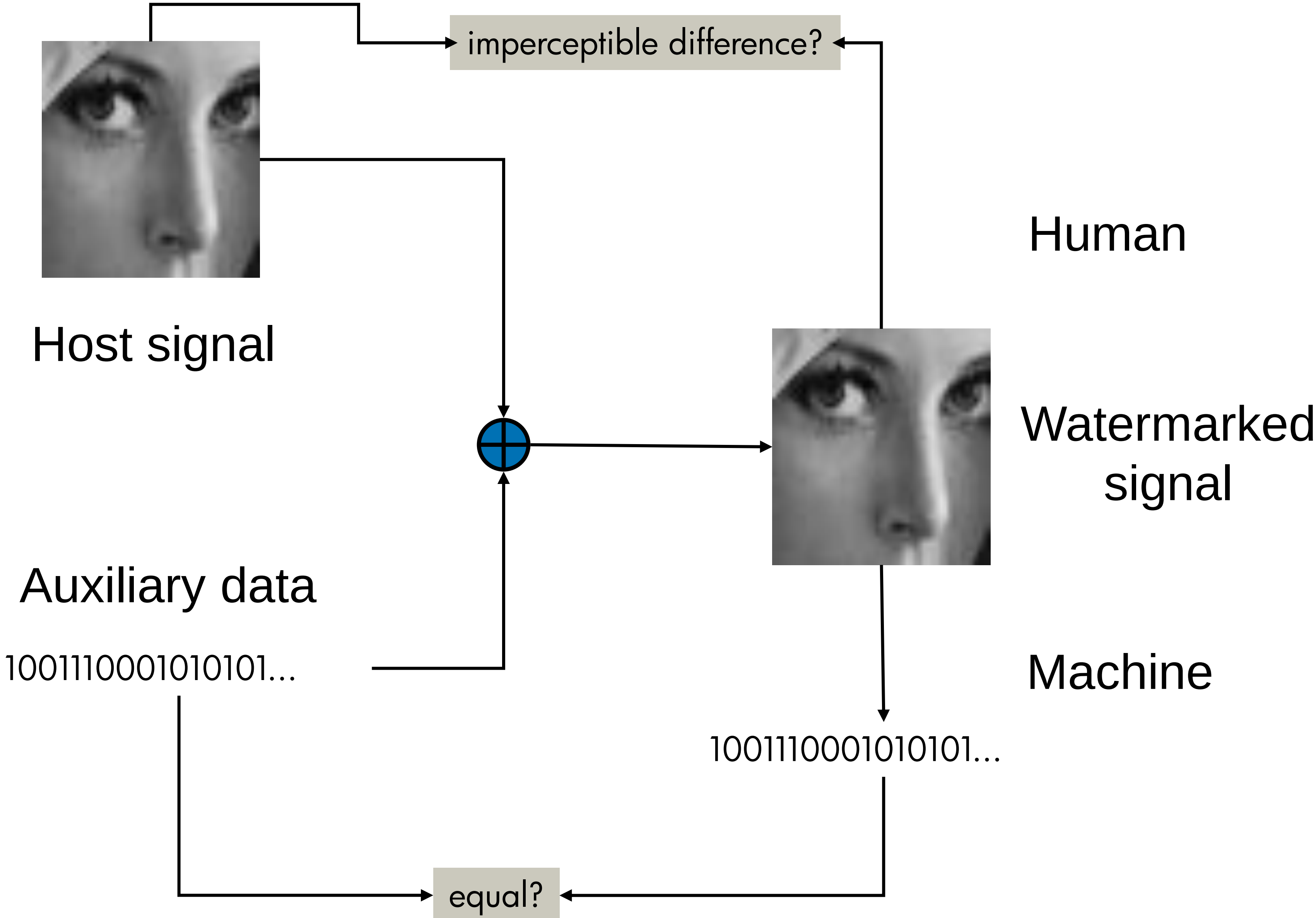
Overview

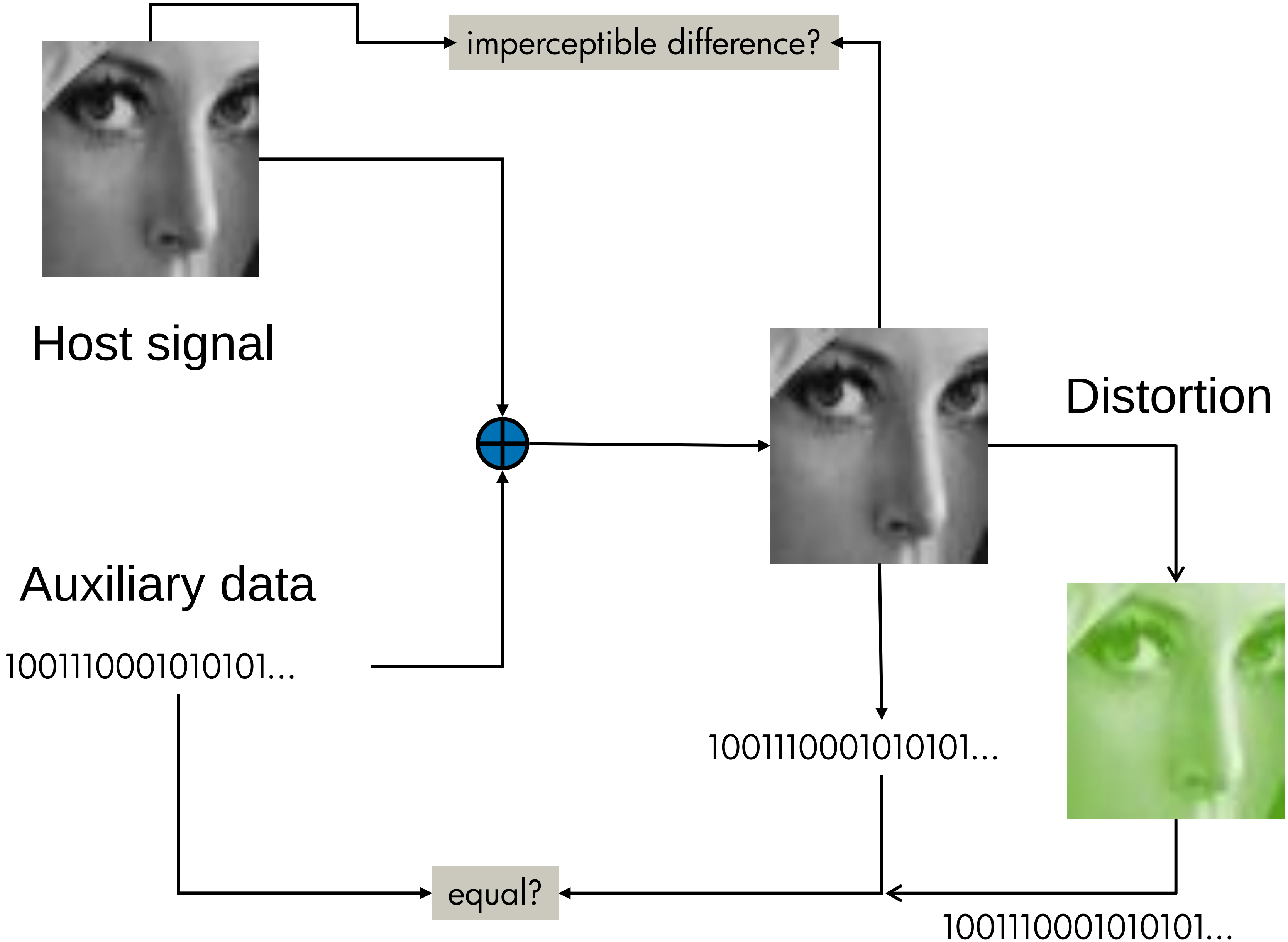
- Definitions
- Applications
- Properties
- Methods
- Security
- Future
- Conclusions

Definitions

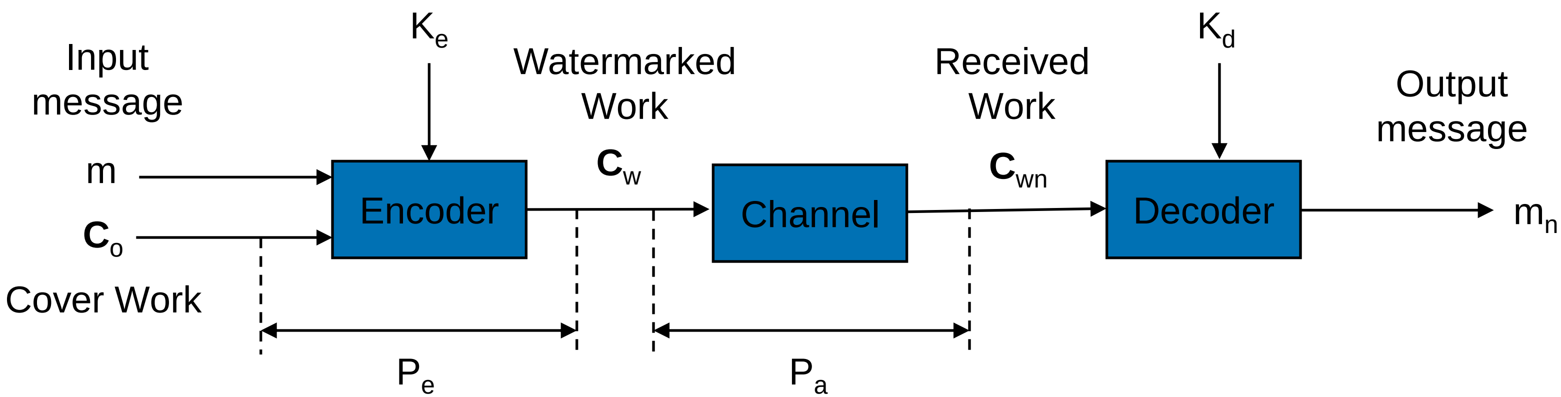
Digital Watermarking

- Original signal
  - host (cover)
    - audio, image, video, 3D model, …
- Auxiliary data
  - potentially related to host
- Multiplexed into one signal
  - Watermarked signal
- Two receivers
  - Human receiver
    - signal detector
    - host signal
  - Mechanical receiver
    - watermark detector
    - auxiliary data


imperceptible difference?
Human
Host signal
Watermarked signal
Auxiliary data
1001110001010101...
Machine
1001110001010101...
equal?


imperceptible difference?
Host signal
Distortion
Auxiliary data
1001110001010101...
1001110001010101...
equal?
1001110001010101...

Formal Model



Applications

Classification

- Robust watermarking
  - Copyright protection
    - Broadcast monitoring
    - Tracking & Tracking
  - Copy protection
    - Copy control signaling (DVD)
- (Semi) Fragile watermarking
  - Authentication
    - Tamper detection
- Steganography
  - Message hiding
- Reversible watermarking

2008/09/02
Philips Enhances Its Content Identification Business With Teletrax Content Monitoring Services

2008/05/05
Teletrax Announces Multi-Year Contract Renewal With The NBC Agency

2008/04/29
New Media Measurement Capabilities Gaining Traction Rapidly

Industry news

2008/08/14
Medialink Board Approves Philips Assuming Full Ownership of Teletrax

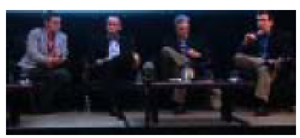
2008/07/31
Digimarc Announces Record Date for Spin-Off

In your industry:

Will mobile devices surpass TV as THE way to view video?

- **Never**
- **By 2015**
- **By 2020**

more



Future of Television West

Teletrax: the way to know when, where and how your content is being used around the world.

How do you: track and monitor video content? Verify broadcast airings: when, where and as contracted? Protect content from piracy or unauthorized re-use? Manage inventory and control video assets?

Teletrax® is the only global media intelligence and broadcast verification service.

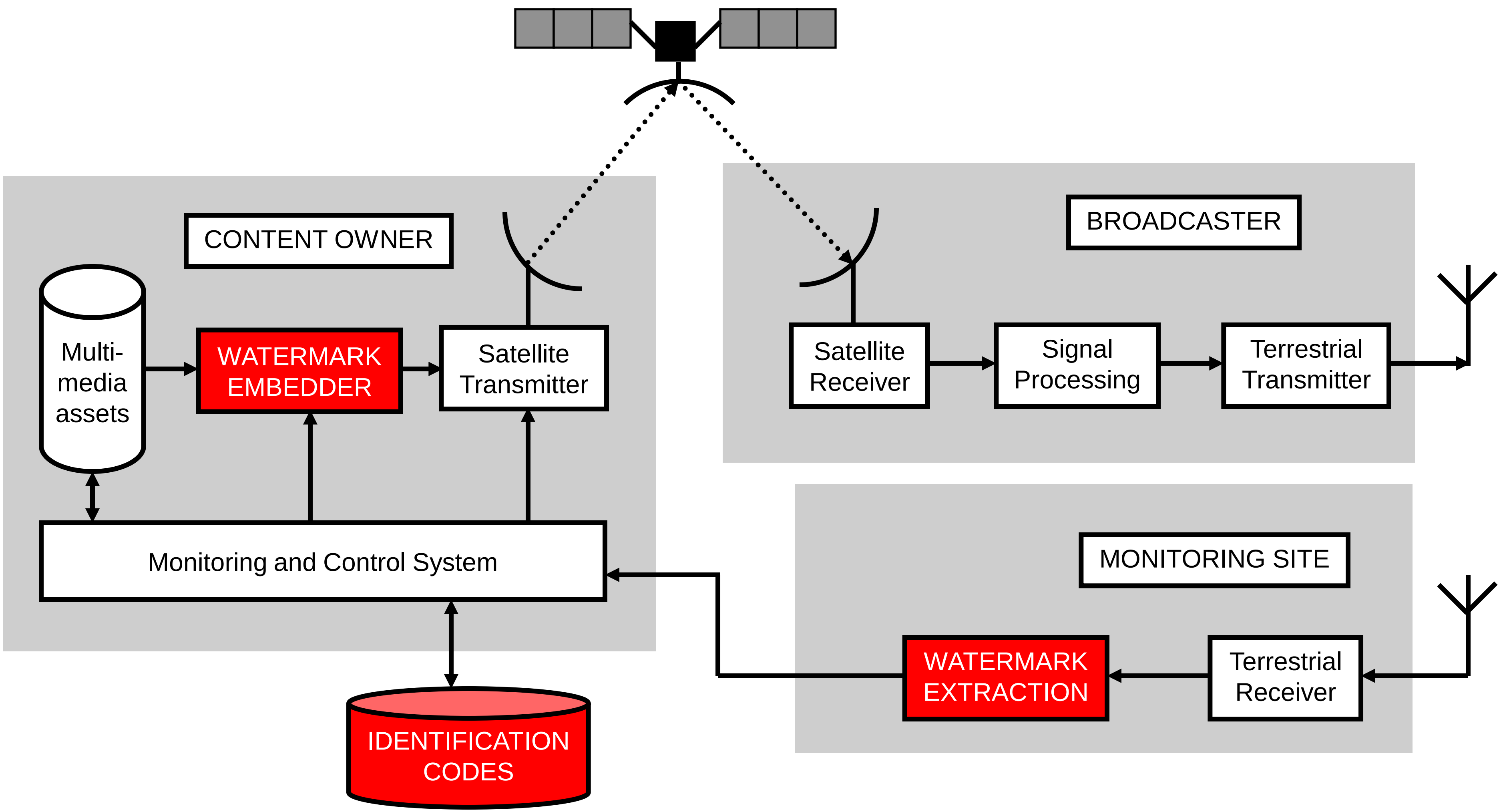
Teletrax enables clients such as **entertainment** studios, **news** and **sport** organizations, **TV syndicators**, **direct response** and **advertisers** to determine precisely **when, where and how** their video content is being used around the world. Over **1,500 television channels** in 50 countries are constantly being monitored, including **all 210** U.S. DMAs.

Teletrax yields critical broadcast intelligence allowing media executives to **keep control** and **better monetize** their assets. A leading broadcast monitoring service provider, Telerax combines cutting-edge **watermarking technology**, from Philips and an intense focus on **customer service.**

For more details **download** the Teletrax brochure, see our **FAQ** or **contact us**.

Every second counts

Broadcast Monitoring

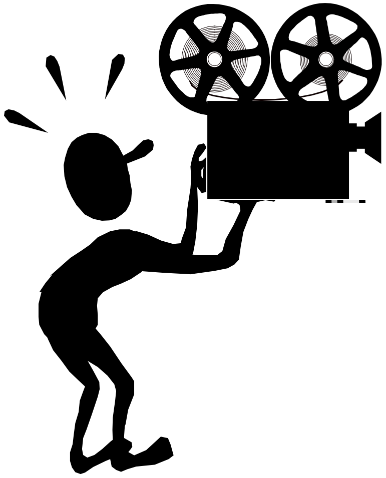


Broadcast Monitoring

- Granularity $\approx$ 1 sec
- Payload $\approx$ ($3\times$) 64 bits
- Robustness
  - MPEG $\geq$ 2 MB/sec
  - PAL-NTSC conversions
  - Change of geometry
- Imperceptible at studio quality
- Affordable monitoring infrastructure
- Security

Classification

- Robust watermarking
  - Copyright protection
    - Broadcast monitoring
    - Tracking & Tracking
  - Copy protection
    - Copy control signaling (DVD)
- (Semi) Fragile watermarking
  - Authentication
    - Tamper detection
- Steganography
  - Message hiding
- Reversible watermarking



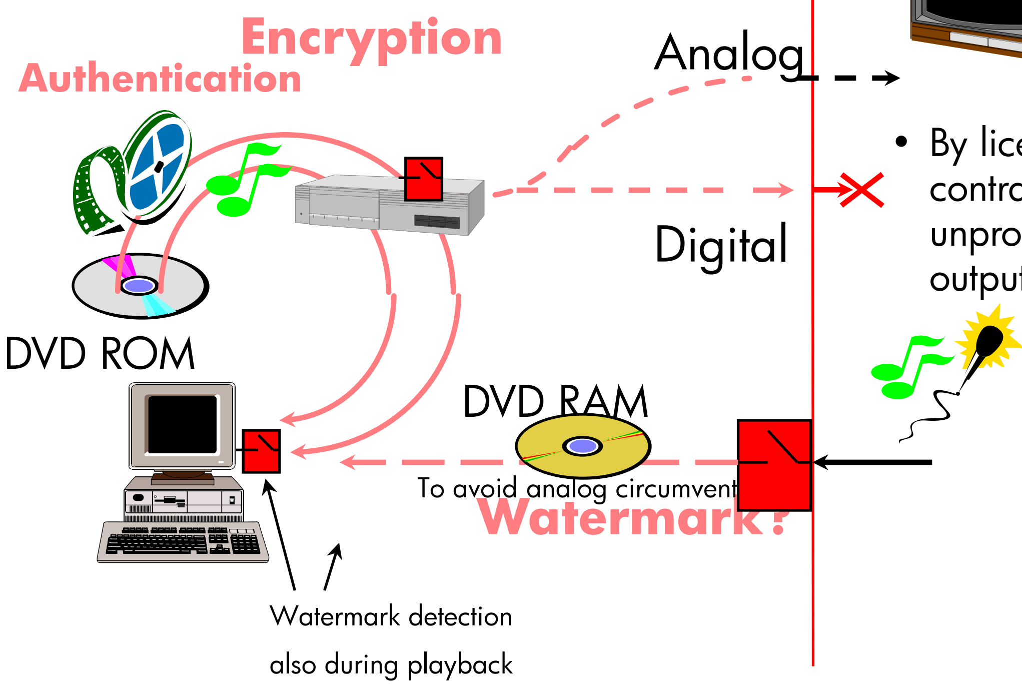
Security

Classification

- Robust watermarking
  - Copyright protection
    - Broadcast monitoring
    - Tracking & Tracking
  - Copy protection
    - Copy control signaling (DVD)
- (Semi) Fragile watermarking
  - Authentication
    - Tamper detection
- Steganography
  - Message hiding
- Reversible watermarking

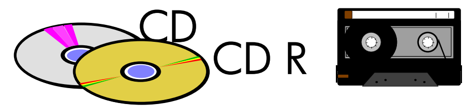
Compliant World

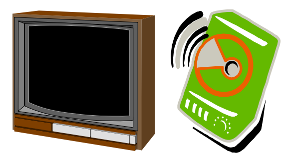
- All content is encrypted on all digital interfaces
- Link-by-link encryption; devices internally process clear content
- Controlled by CSS, 5C, 4C, …
- Includes DVD players, DVD RAM, SDMI audio, DVD audio, PC's



Non-Compliant World

- All analog devices, some digital
- Marginalized by standardization efforts





- Macrovision spoilers
- Watermarks

- By licensing contract no unprotected output
- New laws in US and EU

DVD copy protection

- Four copy states requested
  - Copy-Free (CF)
  - Copy-Never (CN)
  - Copy-Once (CO)
  - Copy-No-More (CM)
- Watermarking for DVD copy protection
  - 1996 – 2000
  - Copy Protection Technical Working Group (CPTWG)
  - Effort failed
    - Security
    - Implementation

Properties

Watermark parameters

- Robustness
  - resistance to (non-malevolent) quality respecting processing
- Perceptibility
  - perceptibility of the watermark in the intended application
- False Positive
  - a positive detection on non-marked content

Watermark parameters

- Granularity
  - minimal spatio-temporal interval for reliable embedding and detection
- Capacity
  - related to payload
  - #bits / sample
- Security
  - vulnerability to intentional attacks
  - Kerkhoffs' principle

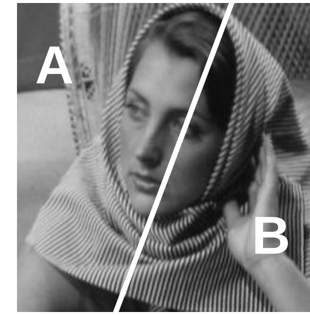
Classification – not discussed

- Robust watermarking
  - Copyright protection
    - Broadcast monitoring
    - Tracking & Tracking
  - Copy protection
    - Copy control signaling (DVD)
- (Semi) Fragile watermarking
  - Authentication
    - Tamper detection
- Steganography
  - Message hiding
- Reversible watermarking

Methods

Spread-Spectrum Watermarking

- 2 disjoint sets, *A* and *B*, of *N/2* pixels each
  - pixels in each set (“patch”) chosen randomly
  - assumption:



$$S = \left(\sum_i A_i - \sum_i B_i\right) / N \approx 0$$

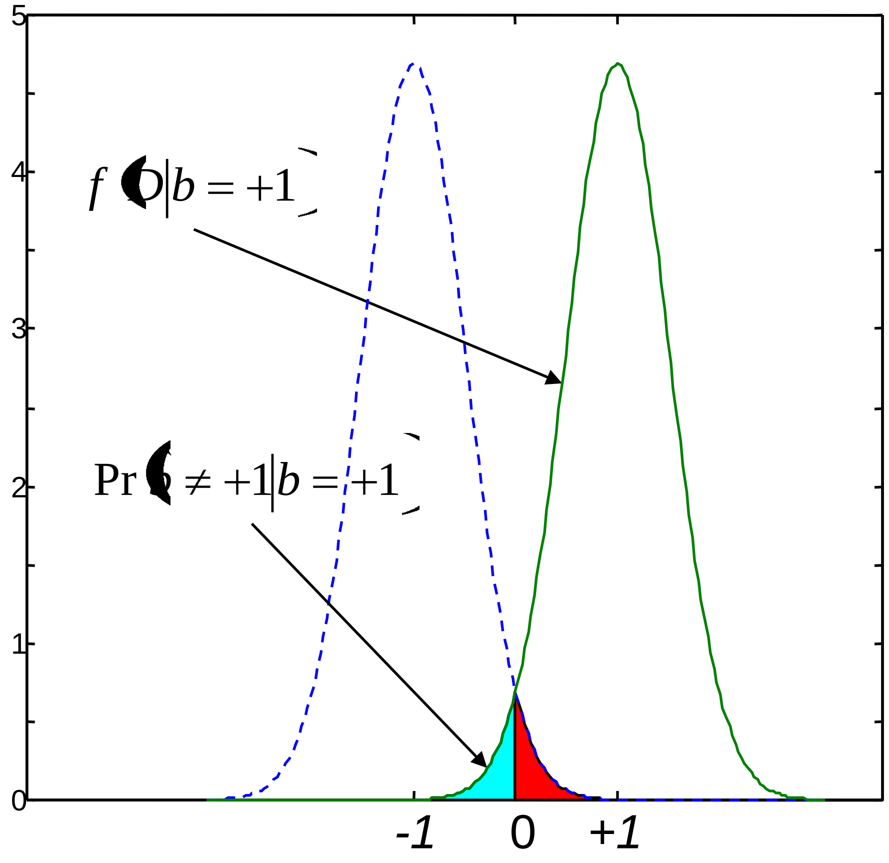
  - embedding bit b ={-1,+1}: $A'_i \leftarrow A_i$+b*1, $B'_i \leftarrow B_i$-b*1

$$S' = \left(\sum_i A'_i - \sum_i B'_i\right) / N =$$
$$\left(\sum_i A_i - \sum_i B_i\right) / N +$$
$$+ (N/2 - (-N/2)) / N \approx b$$

  - if $|S'| \approx 1$, watermark present with value *sign(S’)*
- Prototypical spread-spectrum watermarking
  - communicate information via many small changes

JPEG compression

Additive noise & clipping



5
4
3
2
1
0
$f(\ldots|b=+1)$
$\Pr(\ldots \neq +1|b=+1)$
-1
0
+1

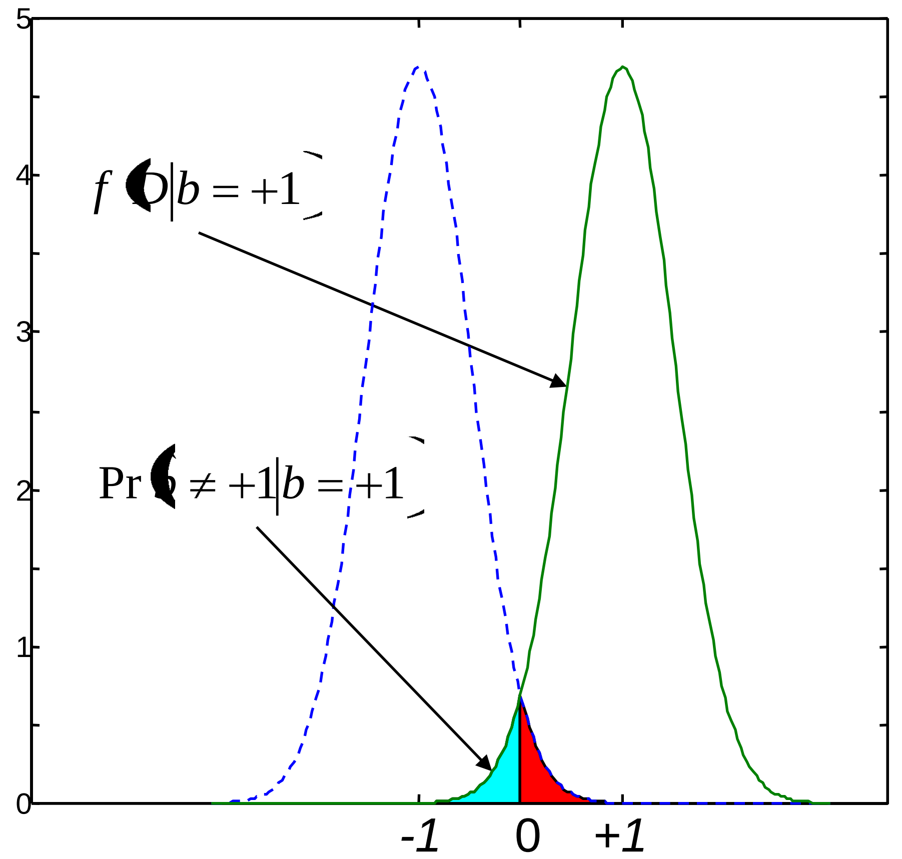
Detection (effectiveness)

- Correlation sum D
  - assumed Gaussian
  - $\sigma_W = 1$
  - variance $\sigma_X^2/(N)$
- Decision rule becomes

$$\hat{b} = \begin{cases} +1, & \text{if } D > 0; \\ -1 & \text{if } D < 0. \end{cases}$$

- Probability of error
  - Q function

$$Q\left(\frac{\sqrt{N}\sigma_W}{\sigma_X}\right)$$



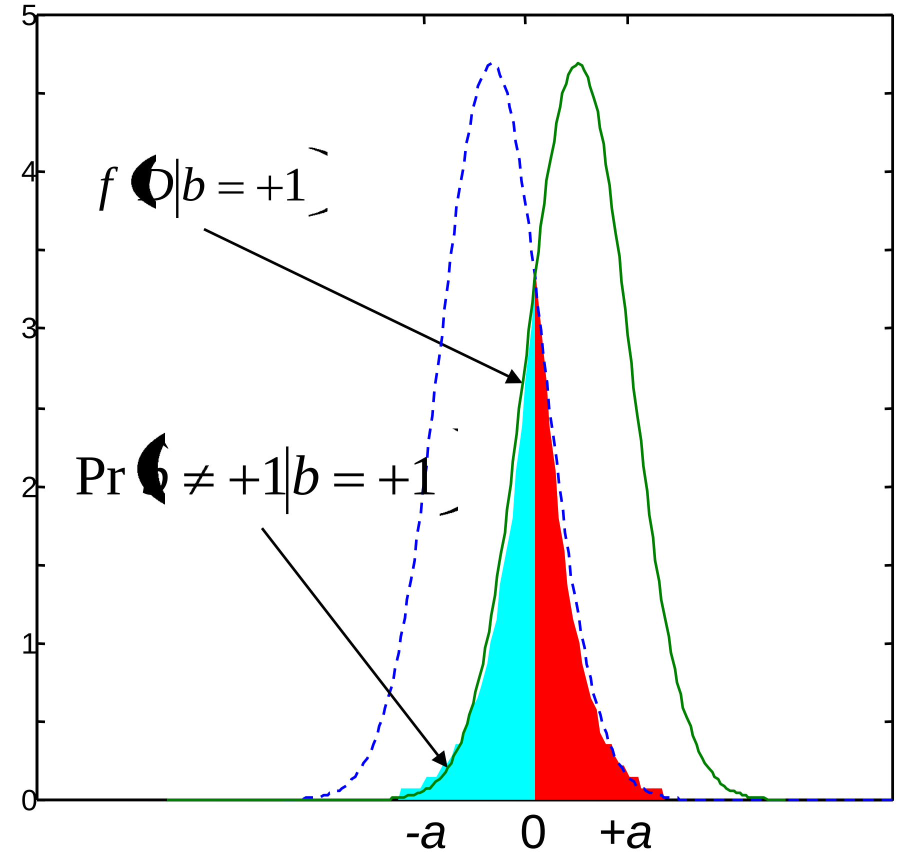
Detection (robustness)

- Correlation sum D
  - assumed Gaussian
  - mean $-a, +a$
  - variance $\sigma_X^2/(N)$
- Decision rule becomes

$$\hat{b} = \begin{cases} +1, & \text{if } D > 0; \\ -1 & \text{if } D < 0. \end{cases}$$

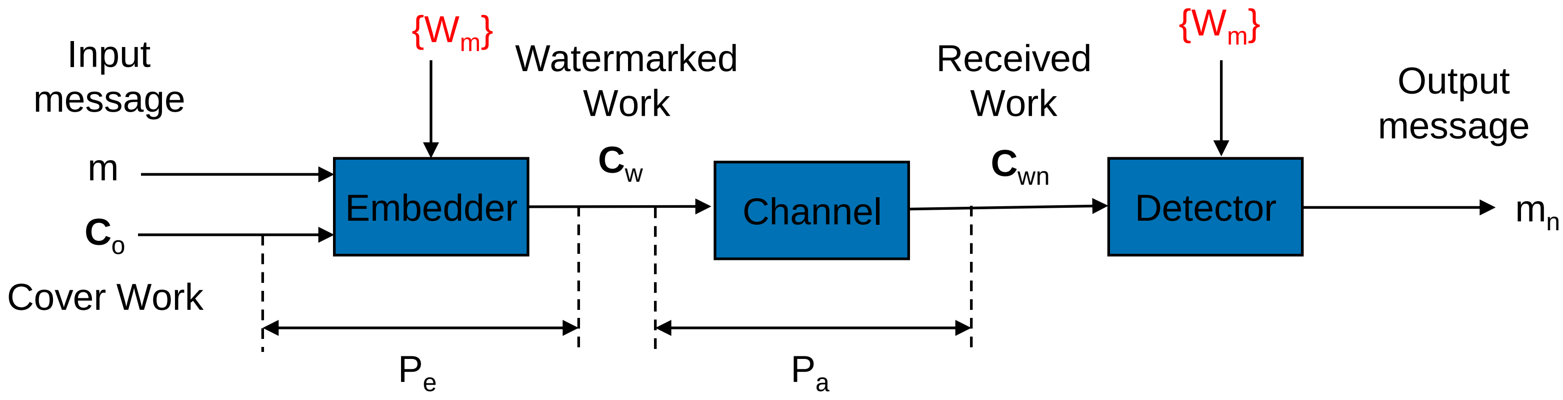
- Probability of error
  - Q function

$$Q\left(a\frac{\sqrt{N}}{\sigma}\right)$$



Transmitting n-bit messages

- Initialization
  - for each message $m \in \{0, \ldots, 2^n\}$ select a watermark sequence W_m
  - Encoder and decoder share the code book $\{W_m\}$
- Loop
  - Encoder chooses message m
  - Encoder adds W_m to host C_o
  - Decoder correlates C_{nw} with every element in code book
  - Decoder declares the message m' such that $W_{m'}$ has the largest correlation with C_{nw}


{Wm}
Input message
m
Co
Cover Work
Embedder
Watermarked Work
Cw
Channel
Received Work
Cwn
{Wm}
Detector
Output message
mn
Pe
Pa

Security

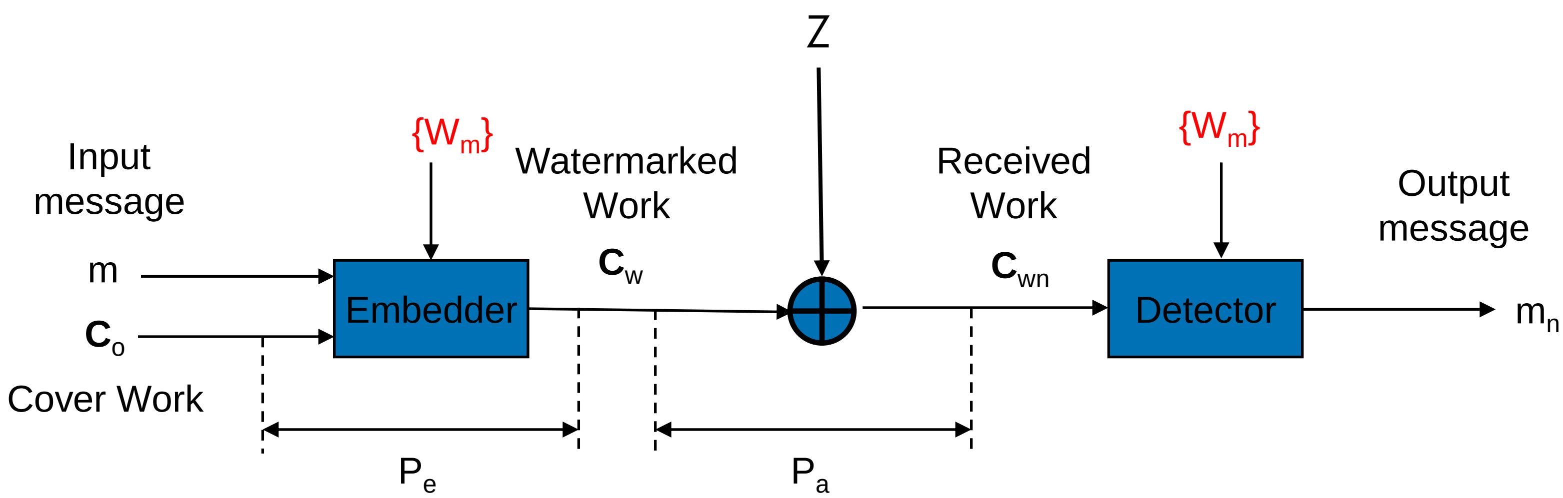
Optimal Rate Question

- Given a some statistical constraints on
  - the host C_o
    - model and energy
  - the embedding distortion P_e
    - type and power
  - the channel distortion P_a
    - type and power
- and allowing for arbitrary long signals,

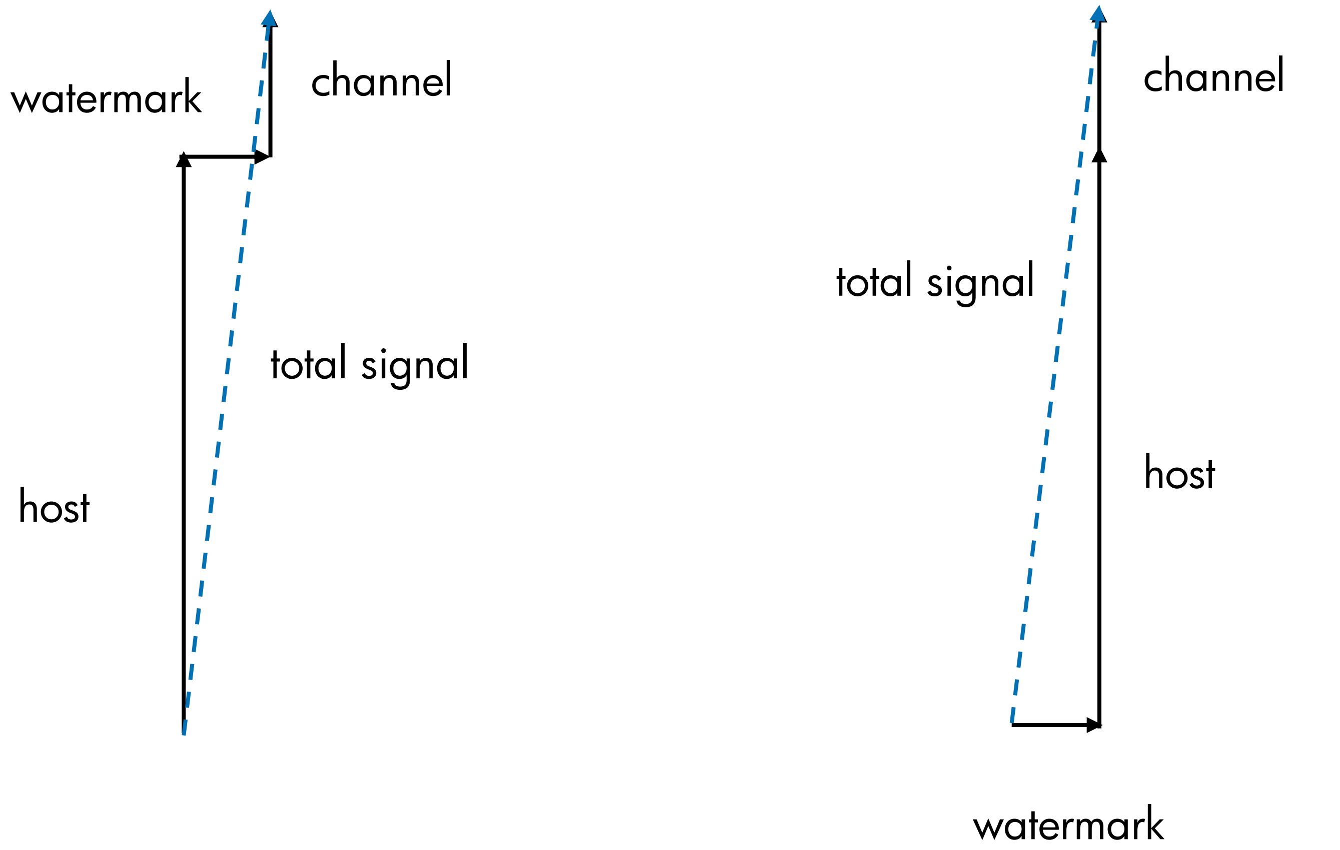
- what is the maximal rate (number of messages per sample) that can be achieved?

Maximal Transmission Rate

- Assumptions
  - C_o is a white Gaussian signal of power P_o
  - The embedding power is restricted to P_e
  - Additive White Gaussian Noise (AWGN) of Power P_a



It is the watermark!



Spread-Spectrum Bound

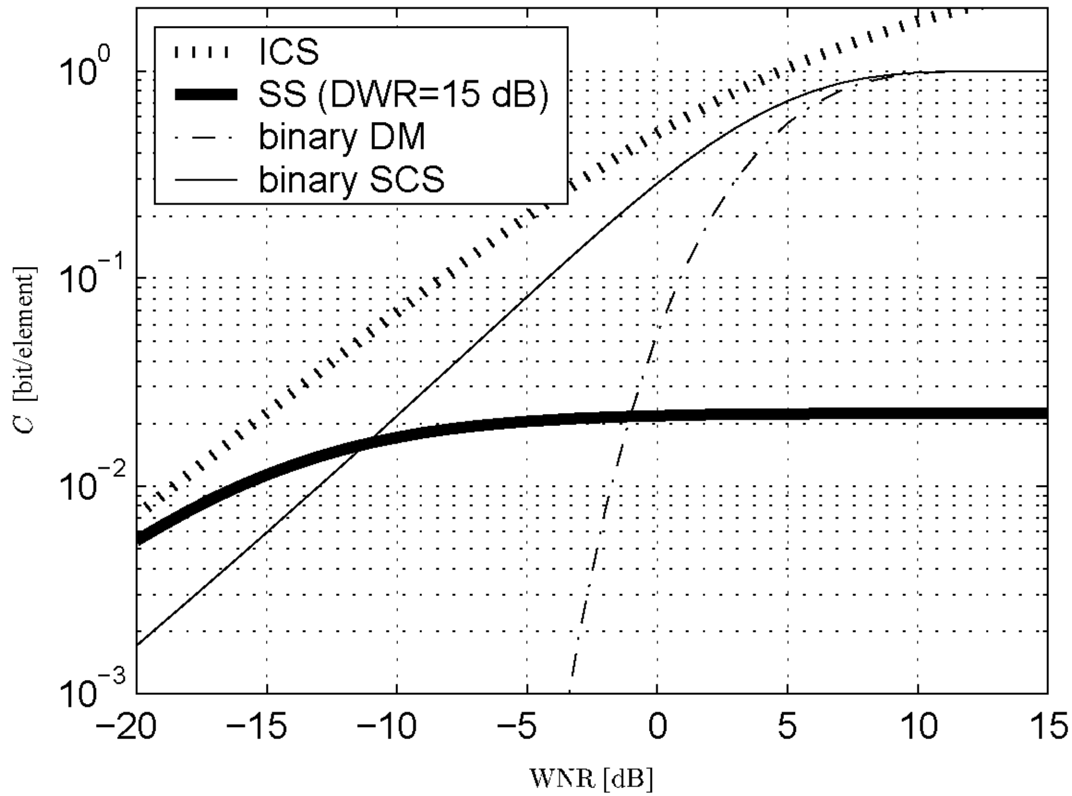
- Observation
  - host signal and channel are AWGN to the watermark signal W_m
- Shannon's Theorem applies

$$R = \frac{1}{2}\log(1 + \frac{P_e}{P_o + P_a})$$

- For small WDR and modest WNR

$$R = \frac{1}{2}\log(1 + \frac{P_e}{P_o})$$

  - Host interference dominates


ICS
SS (DWR=15 dB)
binary DM
binary SCS
10^{0}
10^{-1}
10^{-2}
10^{-3}
−20
−15
−10
−5
0
5
10
15
C [bit/element]
WNR [dB]

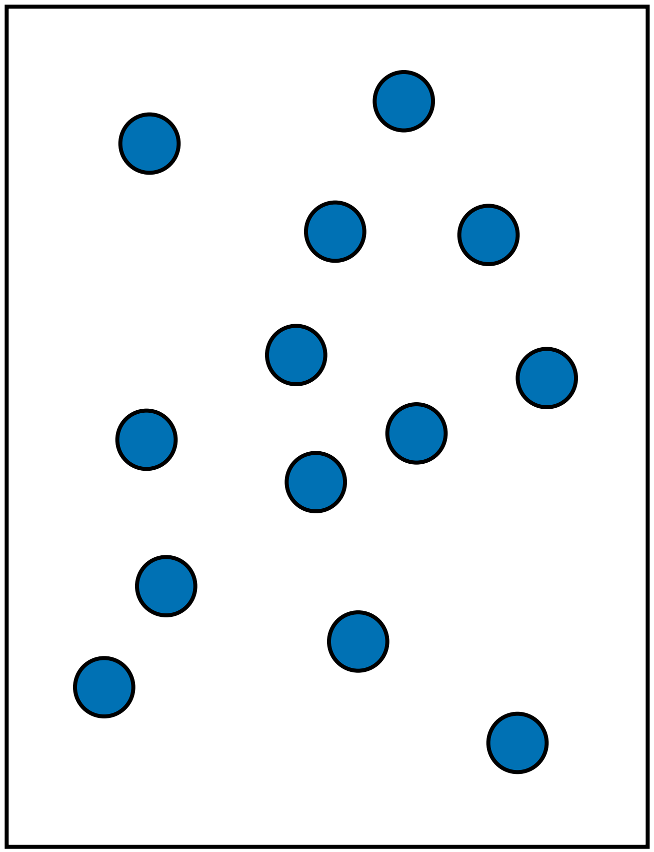
Writing on Dirty Paper

- Shannon
  - Interference only statistically known to the encoder
  - Cause of host interference

- Watermarking
  - Interference known sample by sample to the encoder
  - Potential to do better than Shannon

- Writing on Dirty Paper (Costa, 1985)
  - Host interference can be completely suppressed

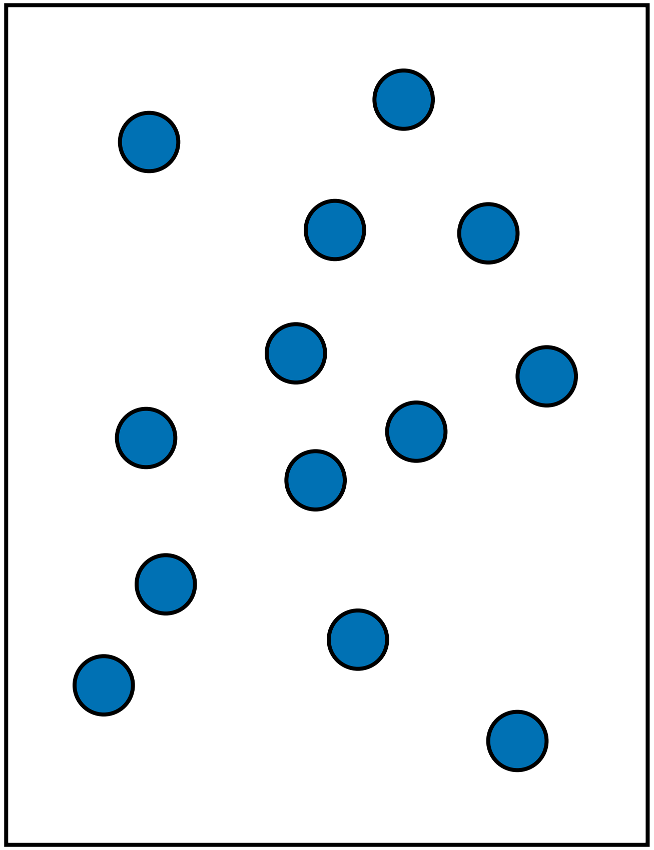
Statistical Knowledge



A

Large distortion

Sample Knowledge



A

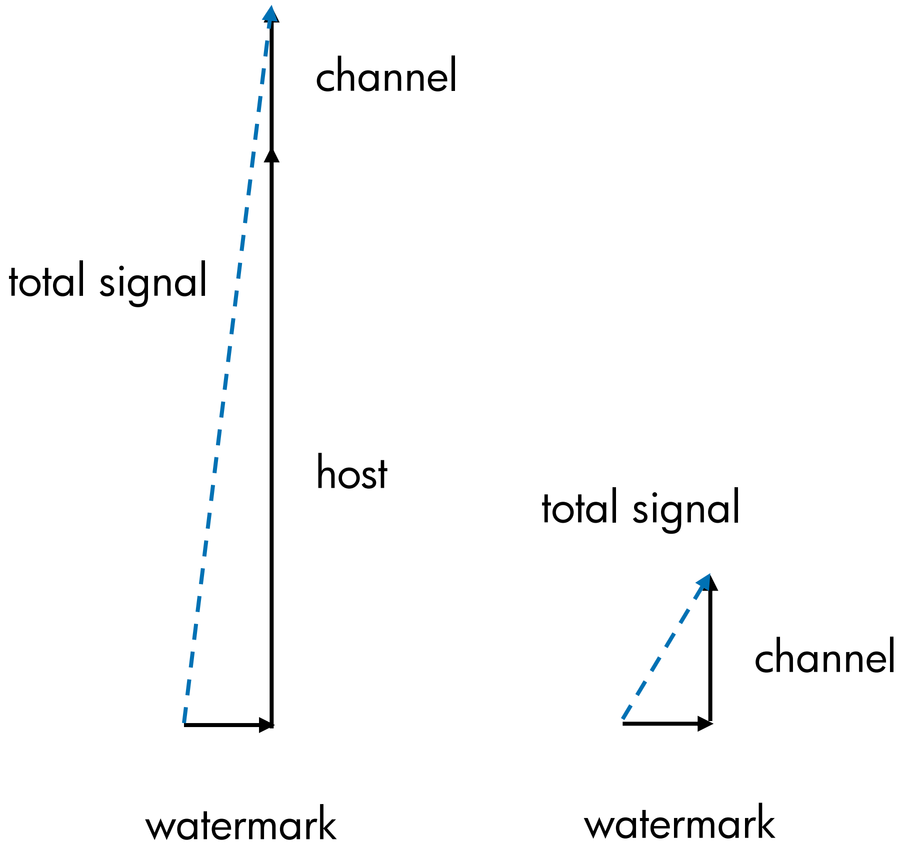
Small distortion

Writing on Dirty Paper

- Host interference can be avoided

$$R = \frac{1}{2}\log(1 + \frac{P_e}{P_a})$$

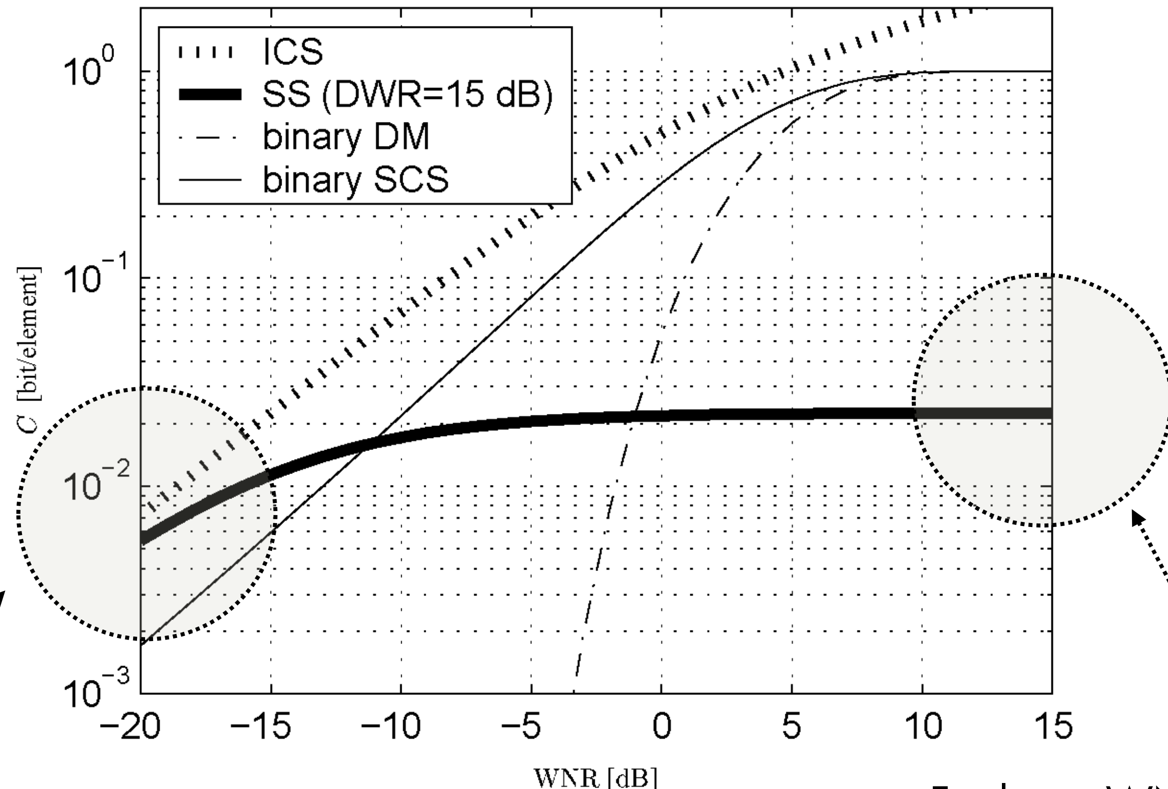
- Associated coding scheme is a variant of LSB watermarking!



invent

Performance graph

Eggers, Girod ©



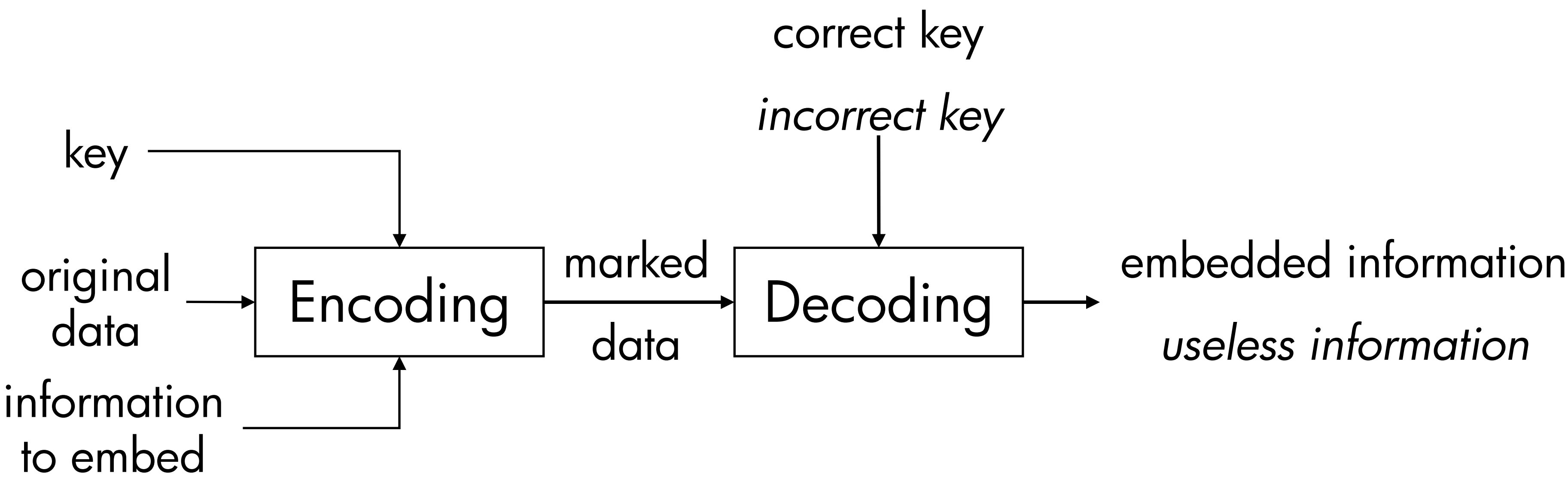
For low WNR Spread-Spectrum approaches rate of optimal scheme ICS

For large WNR Spread-Spectrum underperforms with respect to the ICS scheme.

Security

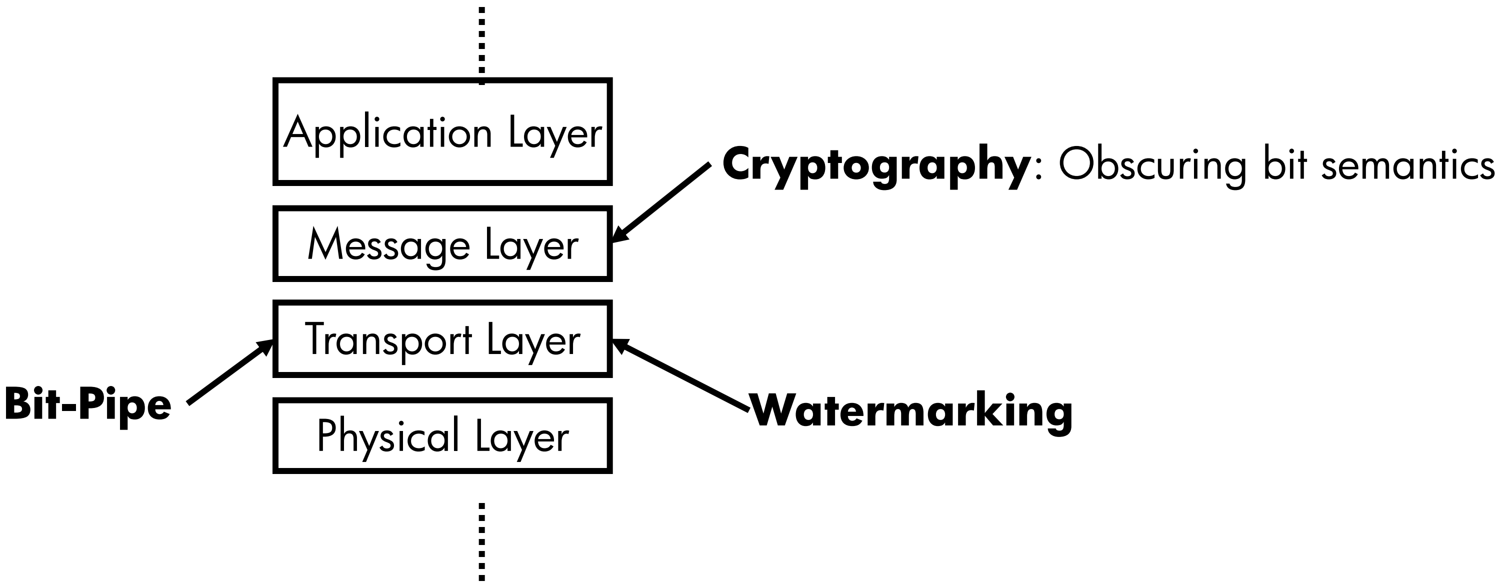
Security - cryptography

- **Definition**: Embedded information cannot be detected, read (interpreted), and/or modified, or deleted by unauthorized parties
- **Kerckhoff's principle**: Security resides in the secrecy of the key, not in the secrecy of the algorithm.

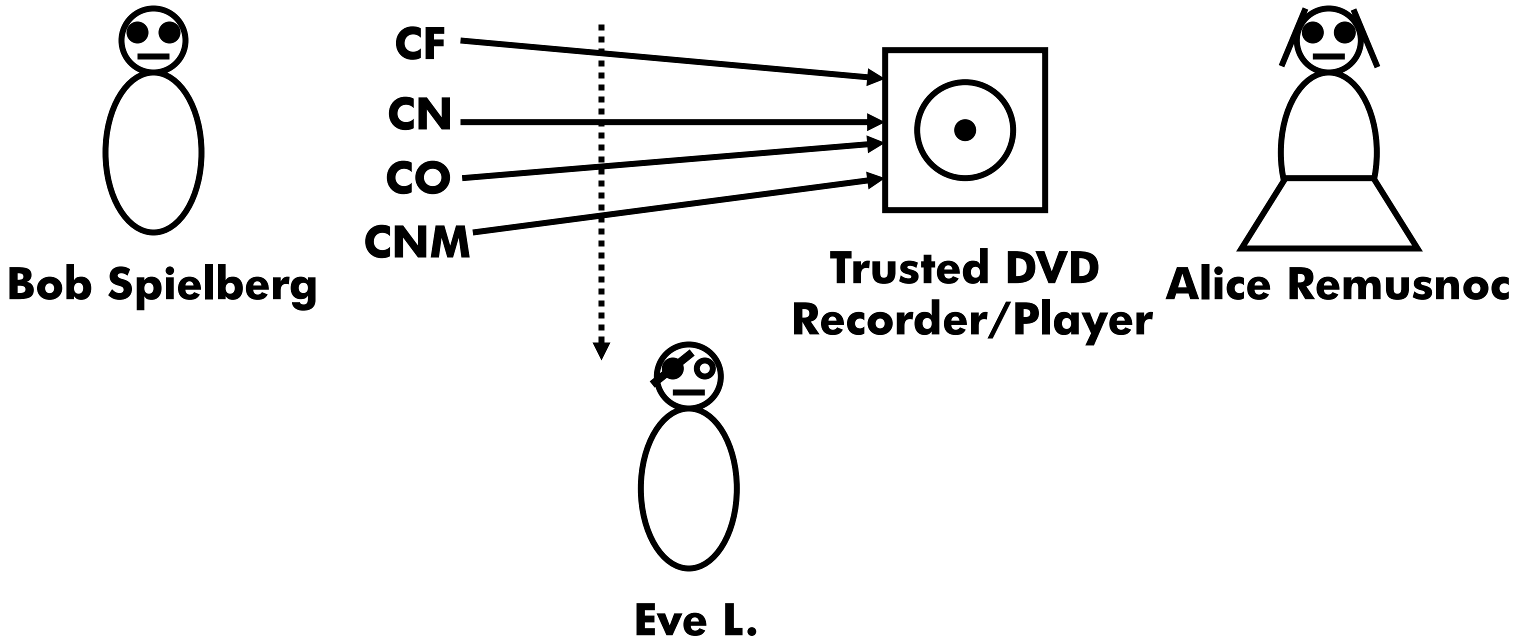


Transmission Security

OSI Terminology



Watermarking Security Example



- Limited message set
- Reading OK
- Alteration NOK!

Lessons Learned

- Watermarking security is <u>not cryptography</u>!

- **Definition**: Watermarking security refers to the inability by unauthorized users to have access to the raw watermarking channel (no semantics).
  - Remove (see previous slide)
  - Read & Estimate
  - Write
  - Modify

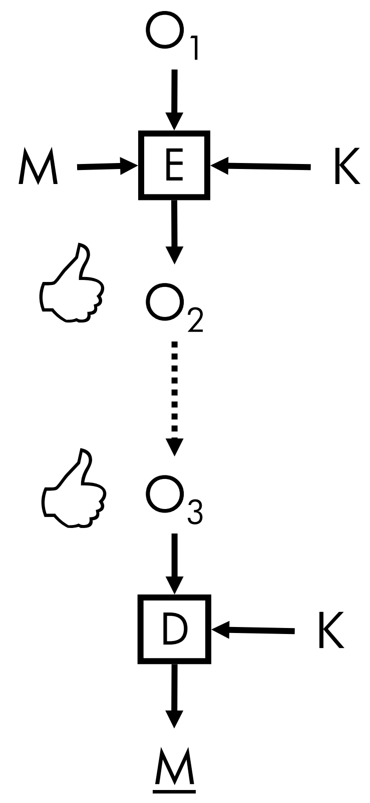
Kerckhoffs' Principle

- Well known cryptographic rule of thumb
  - Security should reside in the keys, not in hiding the algorithm
- Kerckhoffs' principle in watermarking?
  - <u>No</u>: make up for watermarking immaturity
  - <u>Yes</u>: uncertain security is worse than weak security.

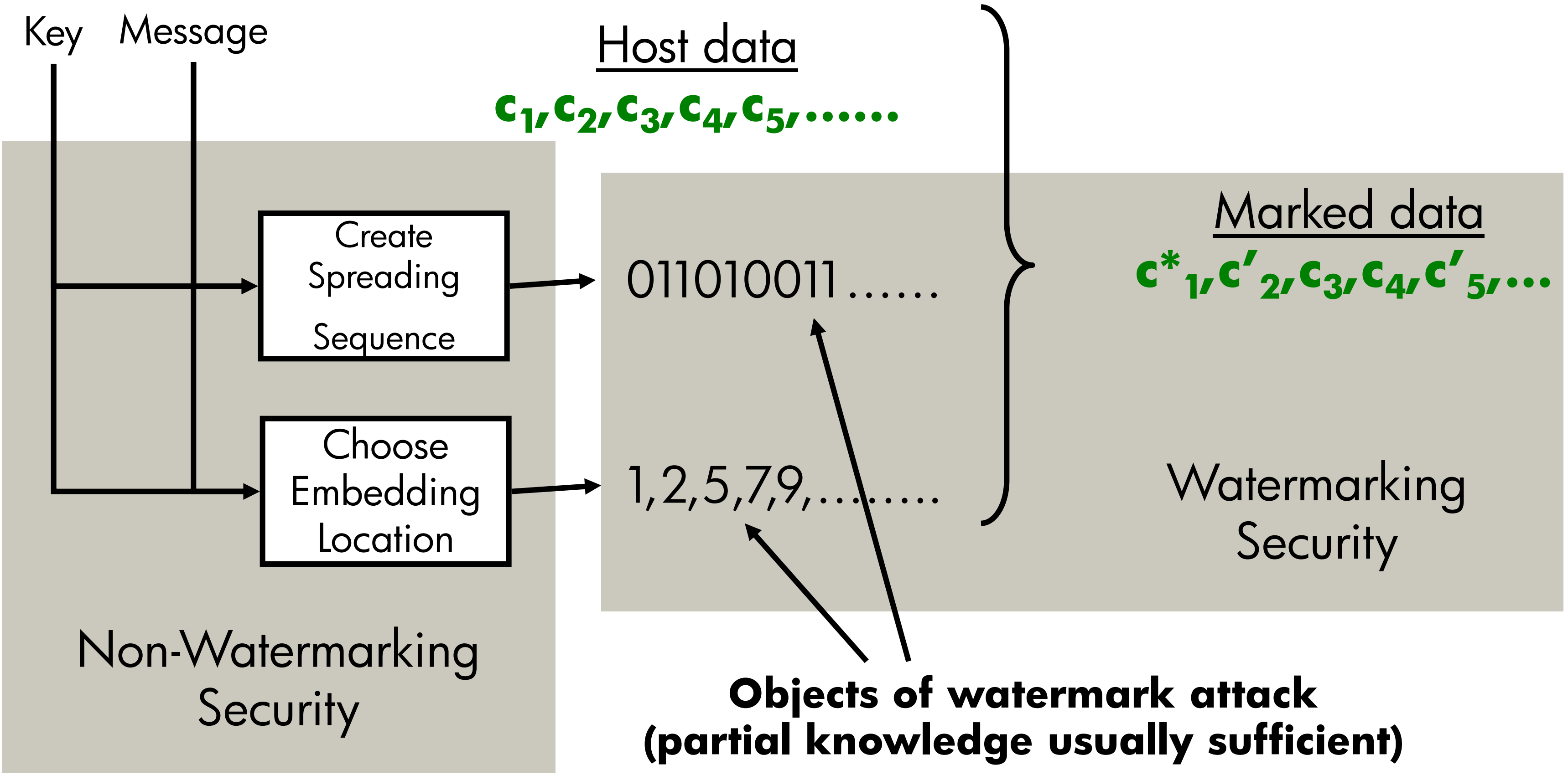
hp invent

Keys in Watermarking

- Key K determines embedding & detection parameters
  - Pseudo-random noise sequences
  - Locations
  - ….
  - In general: internal parameters ‘IP’
- Security of mapping K → ‘IP’ of no relevance to WM security
- Security resides in hiding ‘IP’, i.e. a large key space is needed.



Key Mapping Security

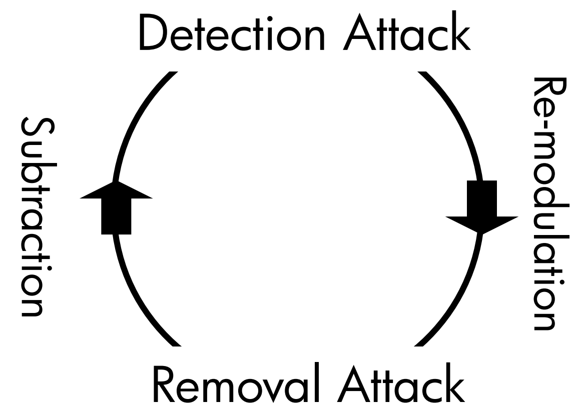


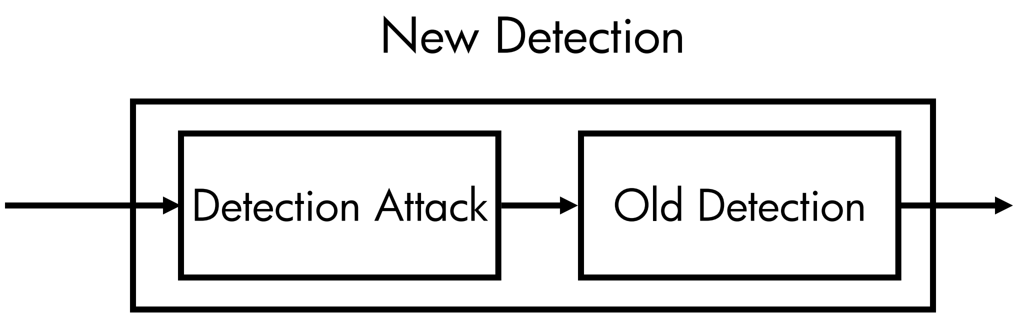
Classification I: Removal Attack

- **Definition**: Unauthorized closing down of the watermark bit-pipe.
- Conflicting classification from literature (Geneva)
  - Removal attack
  - Geometric attack
  - Cryptographic attack
  - Protocol attack
- There is no absolute notion of
  - Geometry
  - Watermark presence
- Methodologically it is better better to merge the first two (Geneva) classes

Classification II: Detection Attack

- **Definition**: Unauthorized observation of the watermark bits (no semantics)
- Strong relation to removal attack
- Detection attacks build better watermark detectors





Classification III: Writing Attack

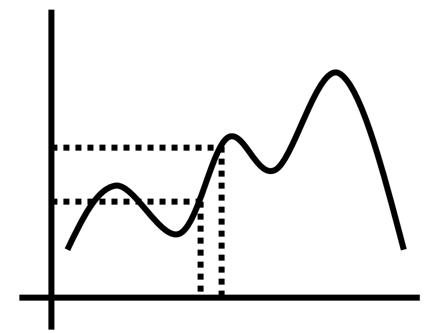
- **Definition**: unauthorized insertion of watermark bits
- Example: Copy Attack (Kutter et al.)
  - **Estimation:**
    - **get hold of a meaningful watermark bits**
  - **Insertion**
    - **Insert bits into other content**
- Countermeasure: content dependent watermarks
  - Every image needs its own random sequence, locations, etc.

Other Attack Parameters

- Number of watermarked objects
  - SOSW, SOMW, MOSW, MOMW
- Availability of tools
  - Embedder, detector
- Knowledge of algorithms: a pointless effort?
  - Limited number of options
  - Literature
  - Patents
- Universality
  - Single objects vs. the complete system

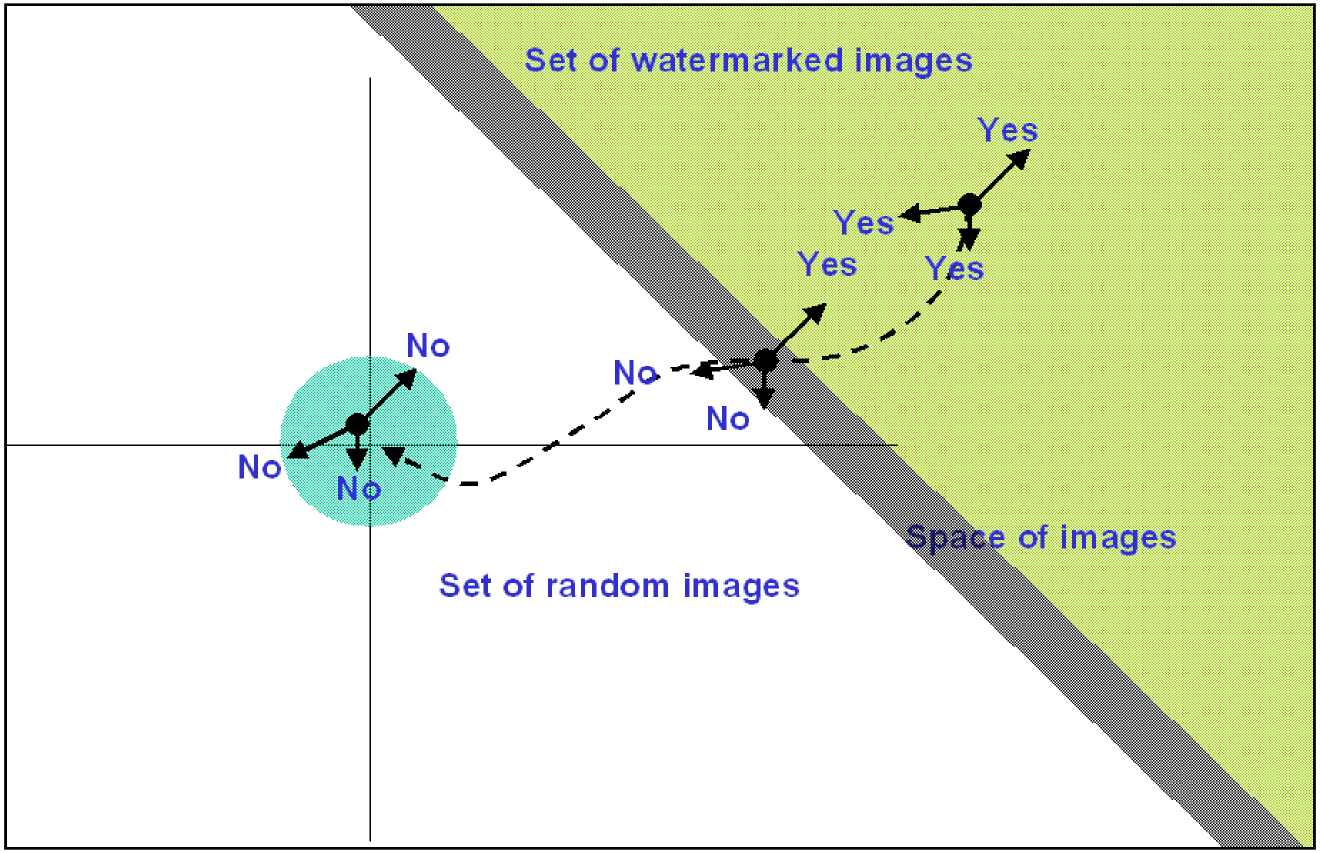
Cryptographic Security of Watermarking?

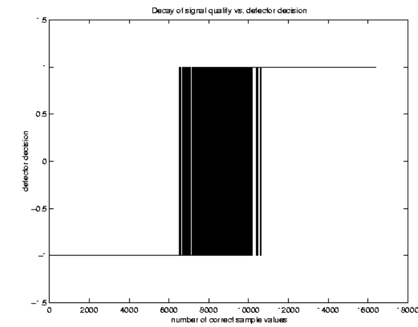
- Cryptography: controlled randomness
  - $M_E = E_K[M]$ is a discontinuous function of M
  - No local invertibility
  - Only way out: exhaustive search



- Robust watermarking: controlled dependency
  - Implied by robustness
  - Continuous dependency on PR sequence, location parameters, etc
  - Local invertibility

Watermark Estimation Through Detector Analysis (ICIP-98)





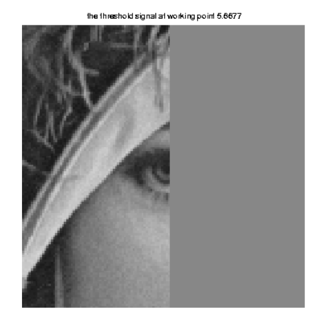


Image at threshold.

Cryptographic Security of Watermarking?

- Unrealistic to expect cryptographic security?
  - SDMI?
  - DVD-Video?
  - Integrated with DRM?
- Maybe not necessary for meaningful deployment
  - Movies/Music/Images are not military-grade secrets
  - Watermarking can provide (commercially) meaningful thresholds
  - Watermarking may not be the weakest link in the chain
- In the end: we always have lawyers to set things straight ☺!

Future

Future

- Watermarking is not a panacea for copy right problems
  - Deployment problems (not discussed)
    - Incentives, costs, authority, legal, …
  - Technical
    - Security of watermarking is not understood
- Challenging problems
  - Modeling watermark security
  - Measuring watermark security
  - Building secure watermark systems

Conclusions

Executive Summary

- Digital watermarking is a well-understood technique for hidden communication.
- However, despite initial high hopes, digital watermarking has not had a major impact on copy right, copy protection and Digital Rights Management.
- One of the main reasons for this state of affair is the poorly understood security aspects of digital watermarking. Improving this understanding is a major open problem.
- Any claim that watermarking is relevant for copyright protection should be taken with a grain of salt.