Digital Watermarking

Chapter I:Introduction
Chapter 2: Applications and Properties

Outline

- Introduction to digital watermarking
  - History and interesting facts
  - Definitions, models, and importance
  - Applications of digital watermarking
  - Properties of digital watermarking

Introduction to Digital Watermarking

What is a watermark?

- Watermarking is an important mechanism applied to physical objects like bills, papers, garment labels, product packing…
- Physical objects can be watermarked using special dyes and inks or during paper manufacturing.

Characteristics of watermarks

- The watermark is hidden from view during normal use, only become visible by adopting a special viewing process.
  - E.g. hold the bill up to light
- The watermark carries information about the object in which it is hidden.
  - E.g. the authenticity of the bill
  - E.g. the trademark of the paper manufacturer

History of watermarking (I)

- The term “watermark” was probably originated from the German term “wassermarke”. Since watermark is of no importance in the creation of the mark, the name is probably given because the marks resemble the effects of water on paper.
- Papers are invented in China over a thousand years ago. However, the first paper watermark did not appear until 1282, in Italy.

History of watermarking (II)

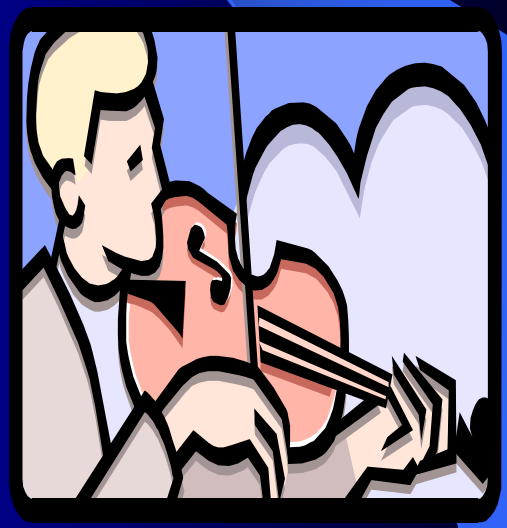
- By the 18th century, watermarks on paper in Europe and America had been used as trademarks, to record the manufactured date, or to indicate the size of original sheets.
- Watermarks are commonly used on bills nowadays to avoid counterfeiting

What is digital watermarking?

- Watermarking can also be applied to digital signals!

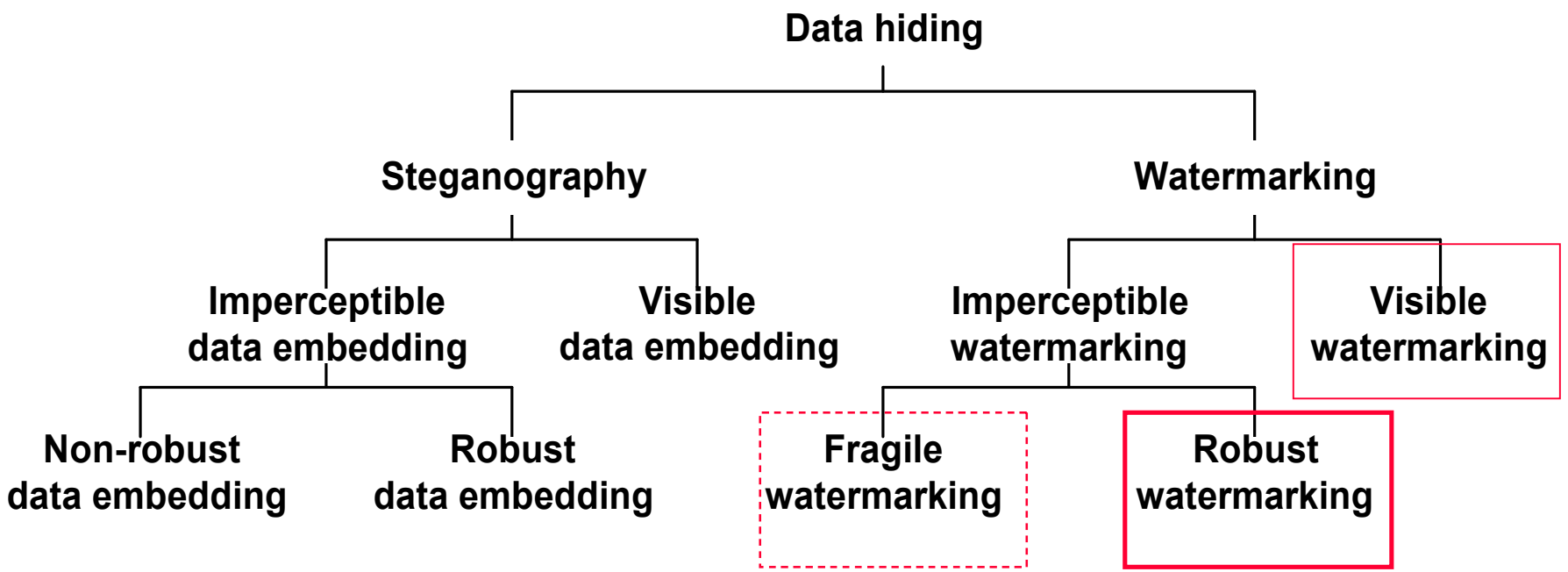
Images

Video



Audio

IPR related information technologies



Information hiding

- Data hiding
  - Containing a large range of problem beyond that of embedding message in content
    - Making the information imperceptible
      - E.g. watermarking
    - Keeping the existence of information secret
      - E.g. anonymous usage of network
      - E.g. hiding portions of database for non-privileged users

Steganography

- A term derived from the Greek words “steganos” and “graphia” (The two words mean “covered” and “writing”, respectively)
  - The art of concealed communication.
  - The very existence of a message is kept secret.
  - E.g. a story from Herodotus
    - Military Messages tatooed on the scalp of a slave

Watermarking v.s. Steganography

- Watermark messages contain information related the cover work
- In steganographic systems, the very existence of the message is kept secret.
  - If the message tatooed on the slave is “the slave belongs to somebody”, then we can regard it as an example of watermarking

Classification of information hiding systems

	Cover Work Dependent Message	*Cover Work Independent Message*
Existence Hidden	Steganographic Watermarking	Covert Communication
Existence Known	Non-Steganographic Watermarking	Overt Embedded Communication

Importance of digital watermarking

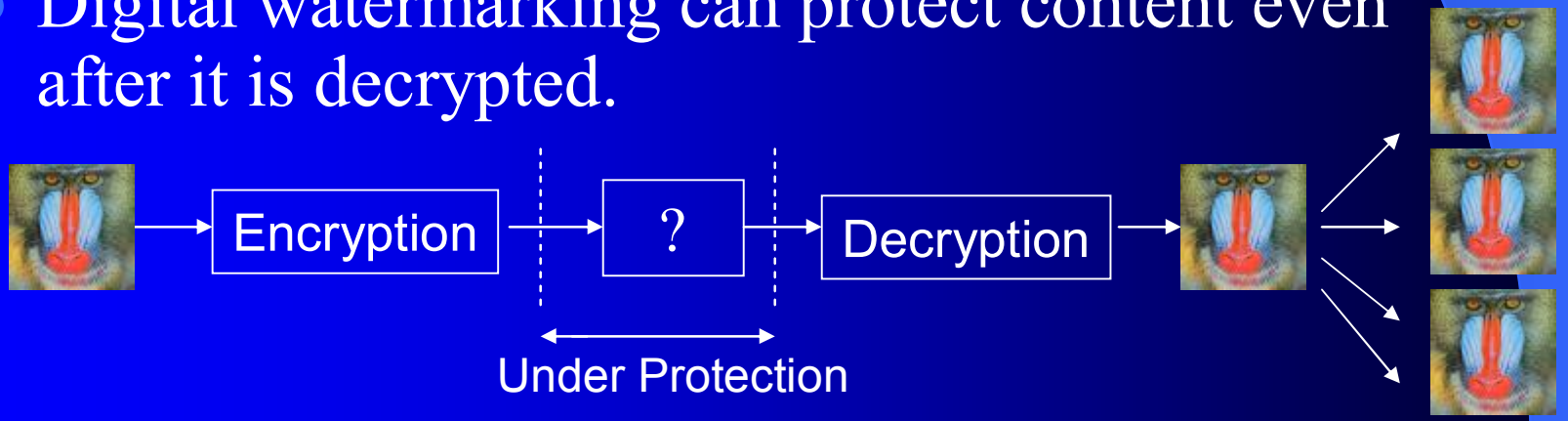
- The sudden increase in watermarking interest is most likely due to the increase in concern over copyright protection of content
- copyright-protected digital contents are easily recorded and distributed due to:

prevalence of high-capacity digital recording devices

the explosive growth in using Internet

Watermarking v.s. cryptography

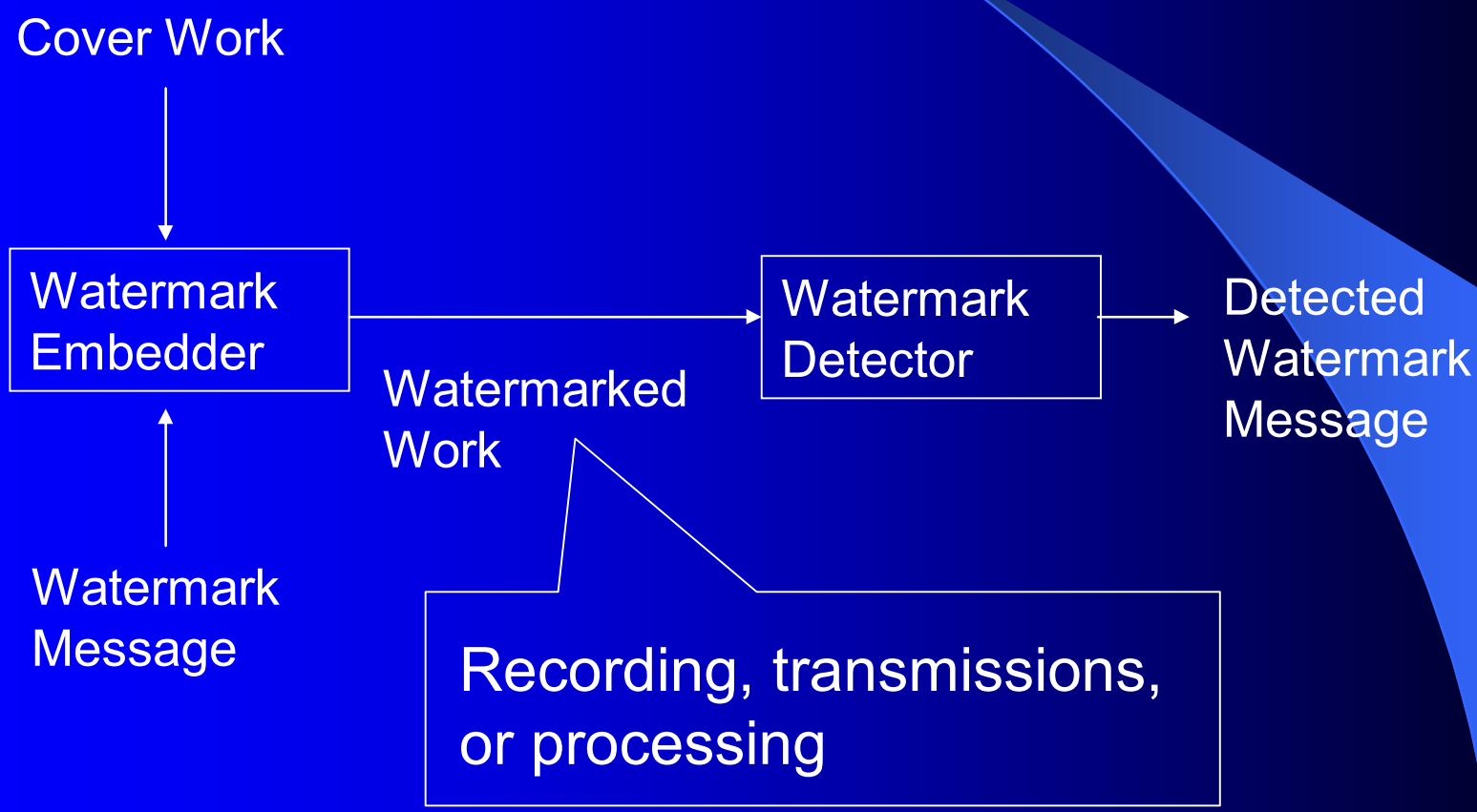
- Cryptography is the most common method of protecting digital content and is one of the best developed science.
- However, encryption cannot help the seller monitor how a legitimate customer handles the content after decryption.
- Digital watermarking can protect content even after it is decrypted.



Definitions about digital watermarking

- Digital watermarking:
  - The practice of imperceptually alternating a Work to embed a message about the Work.
  - Related terms
    - Work: a specific copy of some electronic signal, such as a song, a video sequence, or a picture
    - Cover Work: the original un-watermarked work, since it covers (hides) the watermark
    - Watermark: the messages being embedded, indicating some information about the work

A digital watermarking system



History of digital watermarking

- The first watermarking example similar to the digital methods nowadays appeared in 1954. The Muzak Corporation filed a patent for “watermarking” musical Work. An identification Work was inserted in music by intermittently applying a narrow notch filter centered at 1KHz.
- About 1995, interest in digital watermarking began to mushroom.

Difference between watermarking and other IPR techniques

- Watermarks are imperceptible
- Watermarks are inseparable from the works in which they are embedded
- Watermarks undergo the same transformations as the work

Applications of digital watermarking

- Owner identification
- Proof of ownership
- Broadcast monitoring
- Transaction tracking
- Content authentication
- Copy control
- Device control

Owner identification (I)

- Under the U.S. law, although the copyright notice is not required in every distributed copy to protect the rights of copyright holders, the award to the copyright holders whose work is misused will be significantly limited without a copyright notice found on the distributed materials.
- Traditional textual copyright notices
  - "Copyright date owner"
  - "© date owner"
  - "Copr. date owner"

Owner identification (II)

- Disadvantages for textual copyright notices
  - Easily removed from a document when it is copied
    - E.g. the Lena Sjooblom picture (see the next slide)
  - Copyright notices printed on the physical medium are not copied along with the digital content
    - E.g. the Music CD
  - Occupying a portion of the image and aesthetically reducing the value of artworks
- Since watermarks are imperceptible and inseparable from the work, they are obviously superior to textual copyright notices.

The Lena Phenomenon

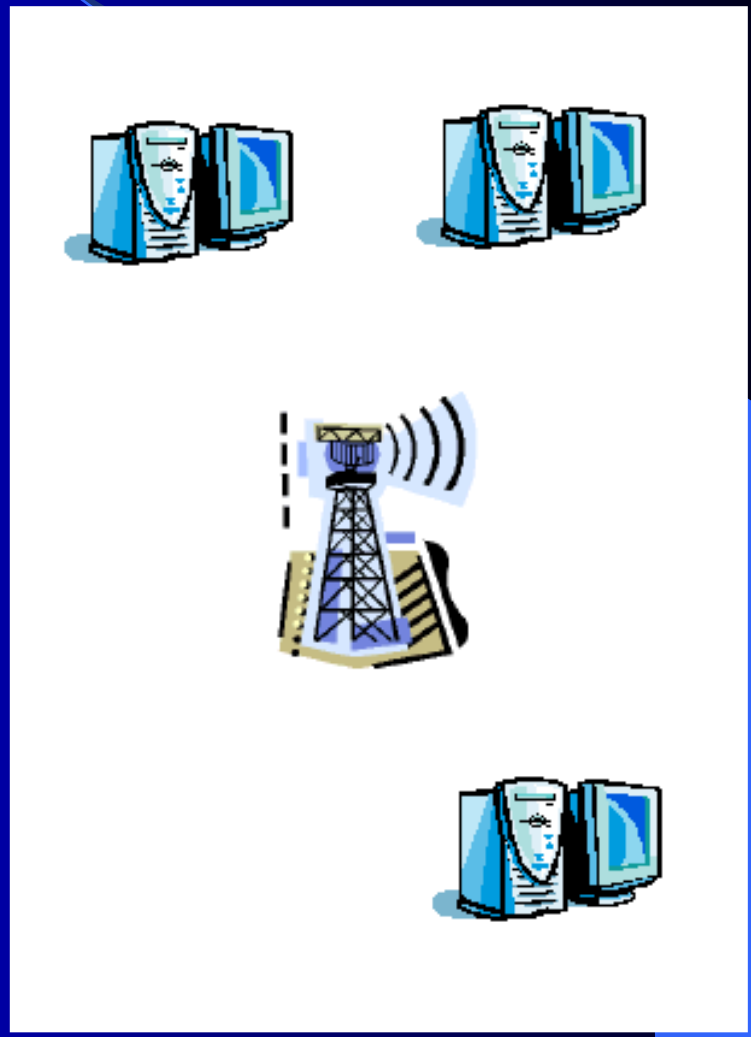
- Lena is the most common test image in image processing research!
- However, the copyright notice of this picture was cropped and ignored.

Proof of ownership

- Textual copyright notices cannot be used to solve the copyright dispute since they can be easily forged
- Registering every work to a central repository is too costly!
  - http://www.loc.gov/copyright
  - $30 per document
- Watermarking can be of use!

Broadcast monitoring (I)

- TV or radio advertisements should be monitored to prevent airtime overbooking!
  - In 1997, a scandal broke out in Japan. Advertisers are paying for thousands of commercials that were never aired!
- Broadcast monitoring
  - By human watchers
  - Passive monitoring
  - Active monitoring



Broadcast monitoring (II)

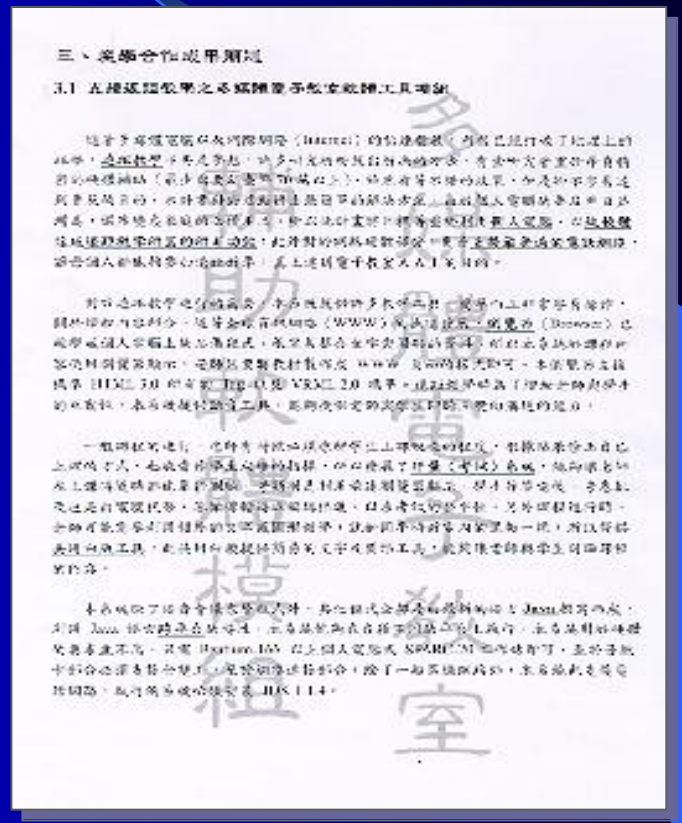
- Passive monitoring
  - Use computers to monitor received signal and compares with a database of known contents
  - Disadvantages
    - Comparing is not trivial
    - Signal degraded due to broadcasting
    - Management and maintenance of the database is quite expensive

Broadcast monitoring (III)

- Active monitoring
  - Simpler to implement
  - Identification information can be directly decoded reliably
  - E.g.
    - close captions on VBI or file headers
  - Watermarking is an obvious alternative method of hiding identification information
    - Existing within the content
    - Completely compatible with the equipments

Transaction tracking

- Watermarks recording the recipient in each legal sale or distribution of the work.
- If the work is misused (leaked to the press or illegally distributed), the owner could find out who is the traitor.
- Visible watermarking is often adopted in this application, but Invisible watermark is even better



The defunct DiVX DVD Player

- The DIVX Corporation sold a enhanced DVD player that implements a pay-per-view model.
- Each player will place a unique watermark in the video disk it played.
- Once the video disk is recorded and sold, the adversary can be tracked!

Copy control (I)

- Encryption is the first and strongest line of defense against illegal copy
  - Overcome an encryption mechanism
    - Decrypt a copy without a valid key
      - Theoretically infeasible for a well designed system
    - Obtain a valid key
      - Reverse-engineering hardware or software
      - E.g. the DeCSS program against the CSS protecting DVD
    - Legally obtain a key and pirate the decrypted content
      - The central weakness of cryptographic protection!
      - The content must be decrypted before it is used, but all protection is lost once decrypted!

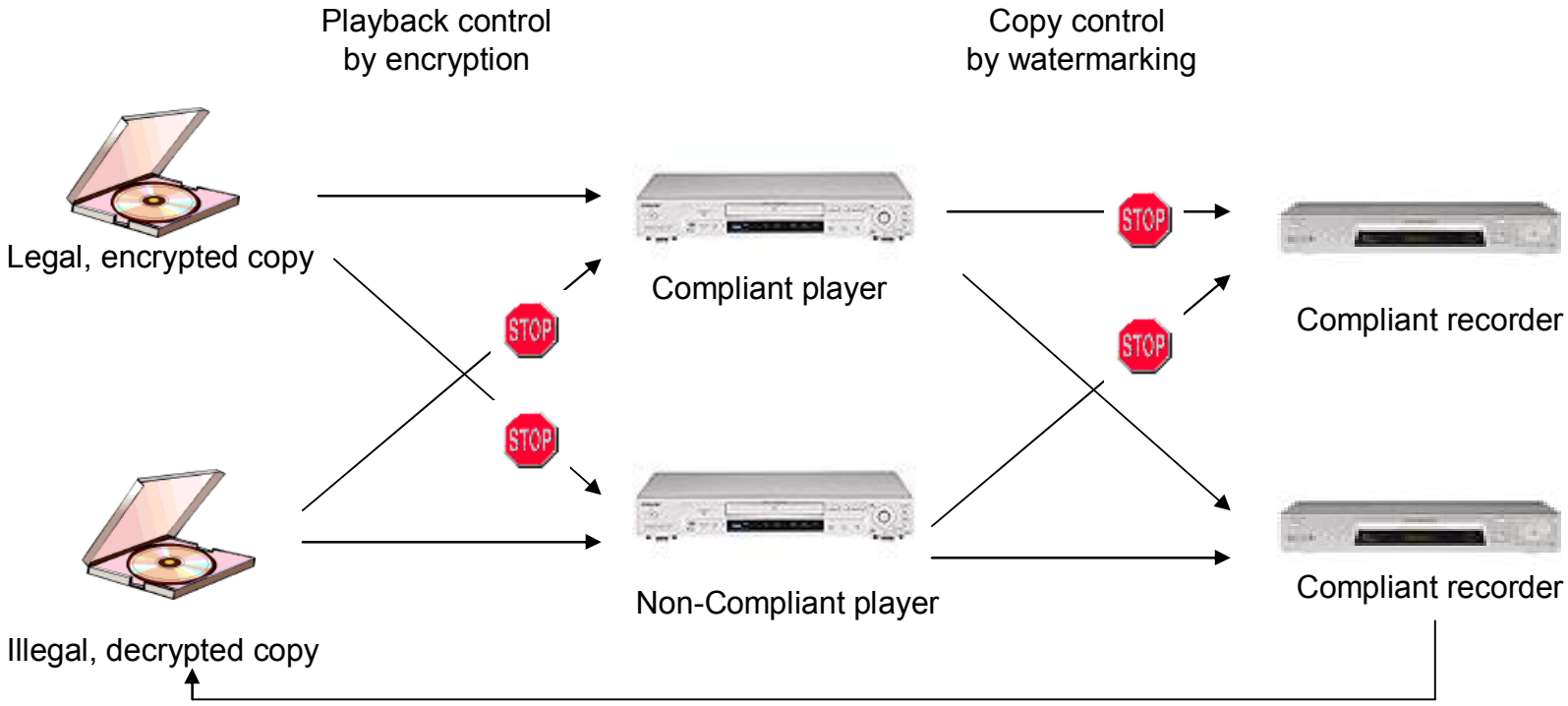
Copy control (II)

- Watermarking in copy control
  - Combining every content recorder with a watermark detector
  - When a copy-prohibit watermark is detected, the recording device will refuse to copy
  - The system has been envisioned by CPTWG and SDMI to protect DVD and audio

Copy control (III)

- Problems of adopting watermarking module in recording devices
  - Increasing cost
  - Reducing the value of devices
- Solution
  - Include the requirement for a watermark detector in the patent license of CSS instead of enforcing by law

Keep honest people honest



Device control

- Copy control belongs to a broader category - device control
- Other applications of device control
  - Automatically turning on/off functions related to special contents
    - E.g Including watermark to skip advertisements
  - Action toys interactive with the TV program
  - Digimarc's MediaBridge

Properties of digital watermarking

- Correct detection result
  - Embedding effectiveness
  - False-alarm rate
- Fidelity (perceptual similarity)
- Resisting distortions
  - Robustness
  - Security
- Data payload (capacity)
- Blind/informed watermarking
- Cost

Effectiveness

- Effectiveness of a watermarking system
  - The probability of detection after embedding
  - A 100% effectiveness is desirable, but it is often not the case due to other conflict requirements, such as perceptual similarity
    - E.g. watermarking system for a stock photo house

False-alarm rate

- Detection of watermark in a work that do not actually contain one
  - The number of false positives occur in a given number of runs of watermark detector
- The false alarm rate of the watermarking system used in DVD recorder should be lower than $1/10^{12}$
  - E.g. a false alarm occurred in a world-series baseball game

Fidelity (perceptual similarity)

- The fidelity of the watermarking system
  - The perceptual similarity between the original and the watermarked version of the cover work
  - It is the similarity at the point at which the watermarked content is provided to the customer that counts
    - E.g. NTSC video or AM radio has different perceptual similarity requirements from the HDTV or DVD video and audio

Problems to determine the fidelity

- Commonly used image similarity index
  - MSE: $\frac{1}{N}\sum_{i=1}^{N}(c[i]-c'[i])^2$
  - SNR: $\frac{\sum_{i=1}^{N}(c[i]-c'[i])^2}{\sum_{i=1}^{N}c[i]^2}$
- Finding a quality index completely reflecting the characteristics of the human perceptual model is difficult

Robustness (I)

- The ability to detect the watermark after common signal processing operations
  - Common images distortions
    - spatial filtering, lossy compression, printing/scanning, geometric distortions
  - Common video distortions
    - Changes in frame rate, recording to tape…
  - Common audio distortions
    - temporal filtering, recording on audio tape…

Robustness (II)

- Not all watermarking applications require robustness to all possible signal processing operations.
- There is a special class of watermarking techniques where robustness is undesirable
  - The fragile watermarking

Security

- The ability to resist hostile attacks
  - Unauthorized removal
    - Eliminating attacks
    - Masking attacks
    - Collusion attacks
  - Unauthorized embedding
    - Embed forgery watermarks into works that should not contain watermarks
    - E.g. fragile watermarks for Authentication
  - Unauthorized detection

Data capacity

- The number of bits a watermarking scheme encodes within a unit of time or within a work.
- Different applications require different data capacities, e.g.
  - 4-8 bits for a 5-minutes video of copy control
  - Longer messages for broadcast monitoring

Blind/informed detection

- Informed watermarking schemes
  - The detector requires access to the un-watermarked original
    - E.g. transaction tracking,
- Blind watermarking schemes
  - Detectors do not require any information related to the original
    - E.g.DVD copy control module
    - E.g. An automatic image IPR checking robot

Multiple watermarks

- In certain cases, more than one watermarks are needed.
  - E.g. American copyright grants the right of TV viewers to make a single copy of broadcast programs for time-shift watch. But further copies is not allowed .
    - Adding two watermarks instead of alternating the original watermark to avoid the risk caused by easily changing watermarks

Cost

- The costs in deploying watermark embedders and detectors depends on the scenario and the business model.
  - Real-time constraint
    - Broadcast monitoring v.s. proof of copyright
  - Embedder/detector constraint
    - Copy protection v.s. transaction tracking (DIV-X)

Watermarking techniques in current standards

- The CPTWG (Copy Protection Technical Working Group) tested watermarking systems for protection of video on DVD disks.
- The SDMI (Secure Digital Music Initiative) made watermarking a core component in their system for music protection.
- Two projects sponsored by the European Union, VIVA and Talisman, tested watermarking for broadcast monitoring.
- The ISO (International Organization for Standardization) took an interest in the context of designing advanced MPEG standards. (MPEG-21)

Companies with watermarking products

- Digimarc bundled its watermarking system with Adobe's Photoshop
- Technology from the Verance Corporation was adopted into the first phase of SDMI and used by some Internet music distributors