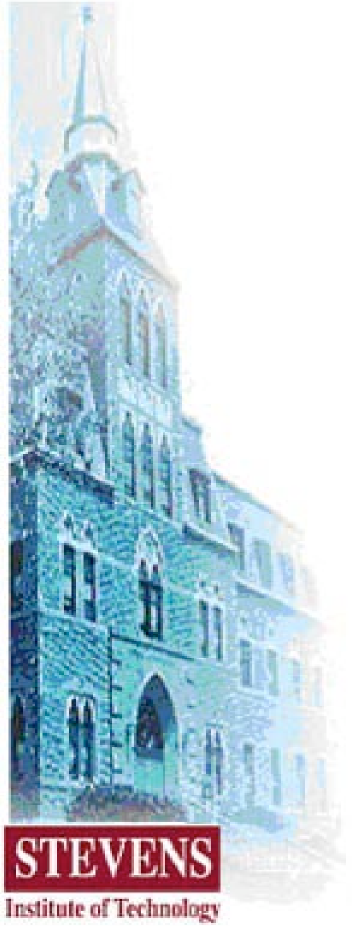


Robust Information Hiding Methods

Outline

- Difference between steganography and watermarking
- Robust data hiding methods
  - Spread spectrum (SS)
  - Quantization Index modulation (QIM)
- Introduction to MPEG/MPEG 4
- A video watermarking example

STEGANOGRAPHY versus WATERMARKING

- **Steganography:**
  - **Hide** a message in ***cover medium*** in such a way that an eavesdropper ***cannot detect*** the presence of the hidden message in the observed **stego medium**
  - May be very sensitive to image processing techniques such as: smoothing, filtering, image compression but should not cause obvious statistical change to disclose the presence of hidden message
  - Often needs a large enough embedding-capacity to communicate the hidden message
  - Original cover medium is not used for extraction

- **Watermarking**:
  - **Hide** a message in ***cover medium*** in such a way that an eavesdropper ***cannot remove or replace*** the hidden message from the **watermarked medium**
  - Need to be very robust to attempts to remove or modify the hidden messag
  - Capacity could be very small
  - Detection may need the cover medium and hidden watermarks. (**Oblivious and non-oblivious)**
    - non-oblivious = original cover medium is needed for extraction
    - oblivious = original covermedium is not necessary

- **Note**: most data hiding methods can be used both for steganography and watermarking

Applications of Steganography

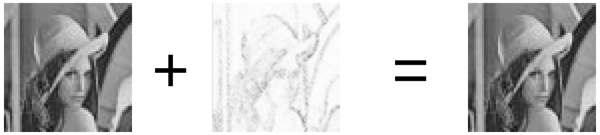
- To have secure secret communications where cryptographic encryption methods are not available.
- To have secure secret communication where strong cryptography is impossible.
- In some cases, for example in military applications, even the knowledge that two parties communicate can be of large importance.
- The health care, and especially medical imaging systems, may very much benefit from information hiding techniques.

Applications of Watermarking

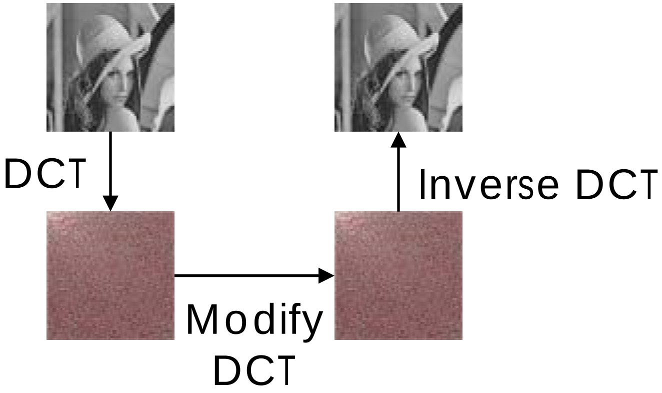
- Robust watermark- Used for copyright protection.
  - Requirements: the watermark should be permanently intact to the host signal, removing the watermark result in destroying the perceptual quality of the signal.
- Fragile watermark- Used for tamper detection or as a digital signature.
  - Requirements: Break very easily under any modification of the host signal.
- Semi Fragile watermark- used for data authentication.
  - Requirements: Robust to some benign modifications, but brake very easily to other attacks.
- Provide information about the location and nature of attack

Spatial Domain and Transform Domain

In the spatial domain, data embedded by directly modifying the pixel values



In the transform domain, data embedded in the transform space by modifying coefficients



The data hiding method in spatial domain is more fragile than that in the transform domain

For color image or video, one of color component for the RGB space is used for embedding. The Luminance component is used for the YCbCr space

Fragile and Robust Embedding

- Fragile
  - Detection fails with even minor modification
  - Useful in tampering detection
  - Common in simple additive stegangraphy
- Robust
  - Detection is accurate even under modification or compression
  - Need for robustness dependent on use of data
  - Often used for digital watermarking

Robust Hiding Schemes

- Spread spectrum – more robust to compression and filtering
  - Add a noise-like signal and detection via correlation
  - Good tradeoff between security, imperceptibility & robustness
  - Limited capacity: host signal often appears as major interferer
- Quantization Index Modulation (QIM) Schemes – data is hidden by choice of quantizer at the encoder- decoder has to decide which quantizer was used to retrieve data
  - Alternative view: switching between two quantizers w/ step size 2Q
  - Robustness achieved by quantization or tolerance zone
  - Trade off between high capacity and limited robustness
- **Note:** both methods are more used for digital watermarking due to their robustness, but they can be used for steganography as well. For simplicity, we only discuss the those embedding schemes used for watermarking only.

Frequency Based Spread Spectrum Watermarking

- Transform image using DCT, DFT, Hadamard, wavelet, key-dependent random transformations
- Select n coefficients to be modified

  - the most perceptually important coefficients

  - fixed band depending on image size

  - key-dependent selection (frequency hopping)
- Generate pseudo-random watermark sequence w_1, …, w_n
- Modulate selected coefficients v_k, $k = 1, \ldots, n$

  $v_k' = v_k + a w_k$, (Ruanaidh et al.)

  $v_k' = v_k + a v_k w_k$, (Cox et al.)

  $v_k' = v_k + a|v_k|w_k$ (Piva et al.)
- Use inverse transform to get the watermarked image

Spread Spectrum in DCT (Embedding)

1000 highest energy DCT coefficients are modulated with a Gaussian random sequence $w_k \in N(0,1)$. The watermark is embedded by modifying the 1000 highest enegy DCT coefficients v_k

$$v_k' = v_k (1 + a w_k),$$

where v_k' are the modified DCT coefficients, and a is the watermark strength also directly influencing watermark visibility.

Spread Spectrum in DCT (Extracting)

Subtract the original image from the watermarked (attacked) image, and extract the watermark sequence η' (may be corrupted due to image distortion)

- Correlate η with η'

η = original watermark sequence

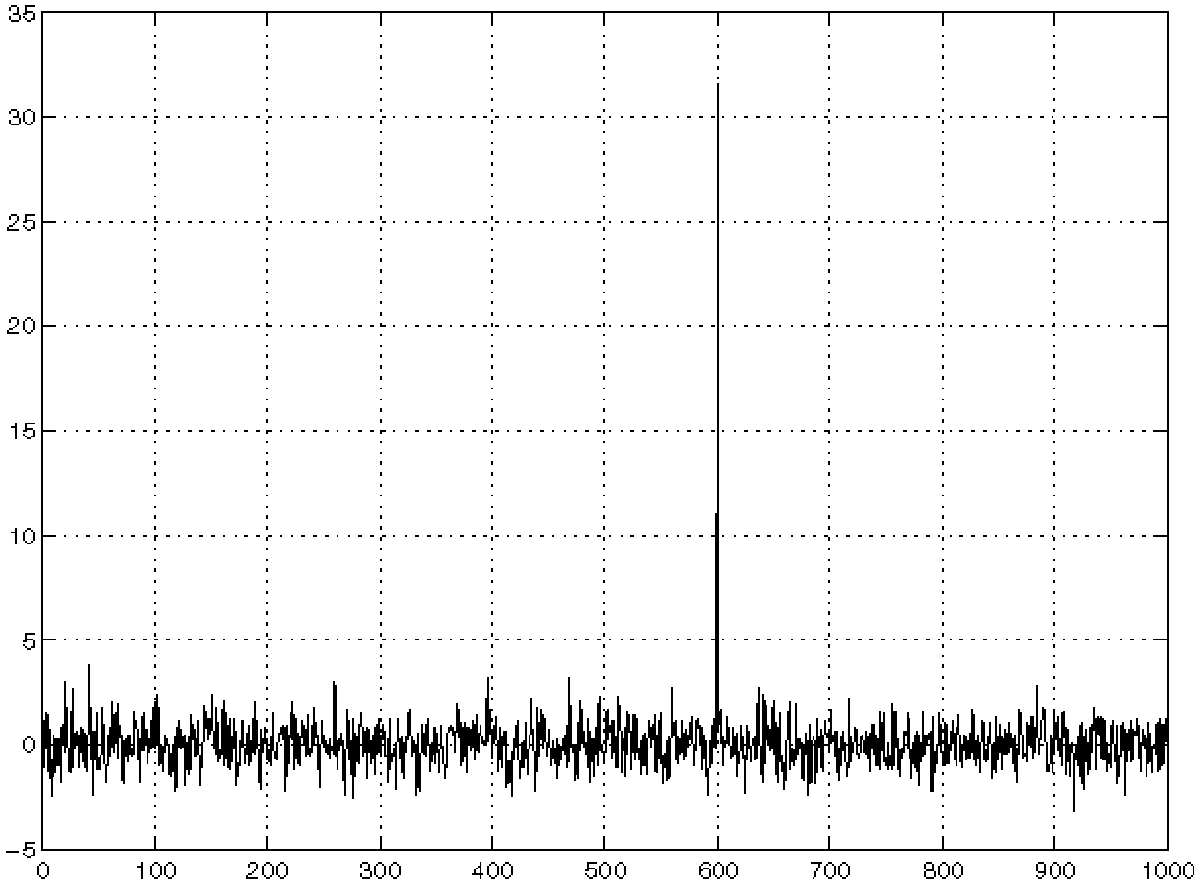
$$\mathrm{sim}(\eta,\eta') = \frac{\eta\cdot\eta'}{\sqrt{\eta'\cdot\eta'}}$$

sim(η, η') is called similarity

sim(η, η') > Th => watermark is present

sim(η, η') < Th => watermark is not present

Watermark detection



Watermark Detection Using Correlation (Non-oblivious)

- Original image v_k
- Watermarked image v'_k
- Attacked watermarked v''_k

$$v_k' = v_k + a w_k, \Rightarrow u_k = (v''_k - v_k)/a$$
$$v_k' = v_k + a v_k w_k, \Rightarrow u_k = (v''_k - v_k)/a v_k$$
$$v_k' = v_k + a|v_k|w_k \Rightarrow u_k = (v''_k - v_k)/a|v_k|$$

- Correlate u_k with w_k
- Threshold the result
- Make a decision about watermark presence

Watermark Detection Using Correlation (Oblivious)

- Correlate v''_k with w_k

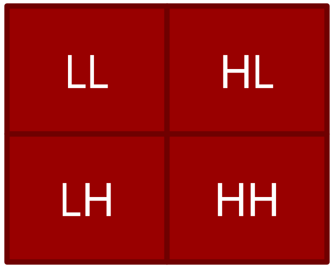
  $v_k' = v_k + aw_k,$
  $v_k' = v_k + a|v_k|w_k$

- If no distortion is present
  $corr = \sum v''_k\, w_k = \sum (v_k + aw_k)w_k \cong an\sigma^2$
  $corr = \sum v''_k\, w_k = \sum (v_k + a\;|v_k|w_k)w_k \cong an\langle|\mathrm{v}|\rangle\sigma^2$

- If incorrect noise sequence is used
  $\langle corr\rangle = 0$ with $\langle \mathrm{corr}^2\rangle \cong n$
  which enables us to set a decision threshold

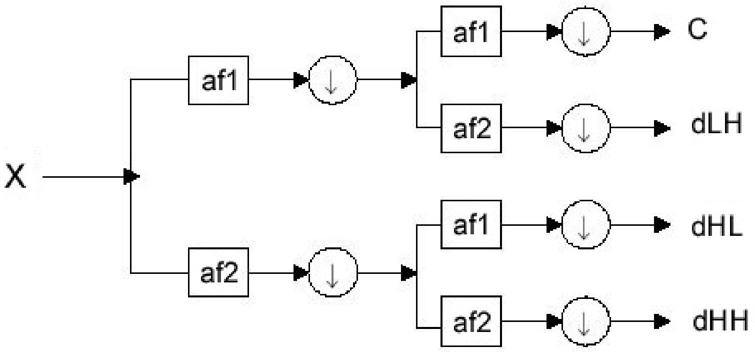
2D – Discrete Wavelet Transform

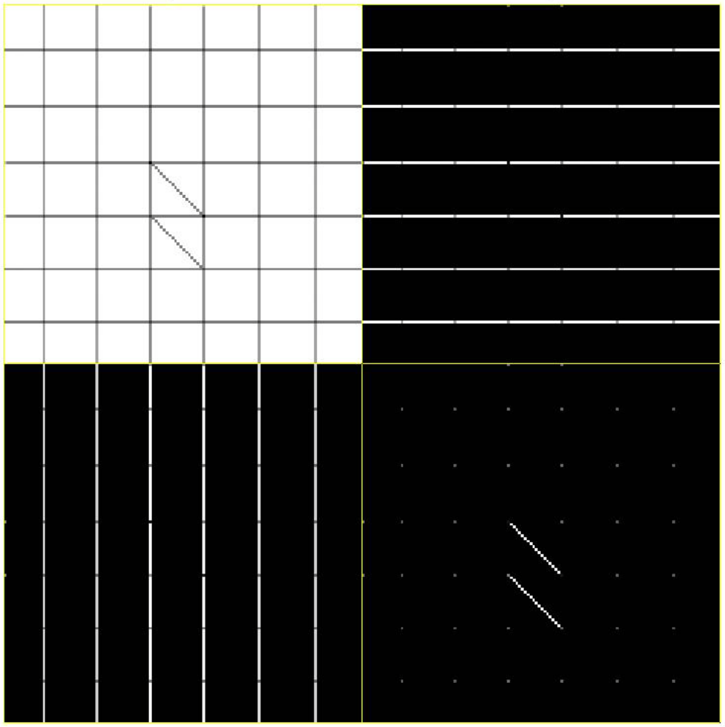
- The 2D-DWT Transform divides the image into 4 sub-bands
  - LL – Lower resolution version of image
  - LH – Horizontal edge data
  - HL – Vertical edge data
  - HH – Diagonal edge data
- Most DWT watermarking algorithms embed only in the HL, LH and HH sub-bands



LL	HL
LH	HH

2D – Discrete Wavelet Transform





Decomposition at level 1

Watermark Embedding Method

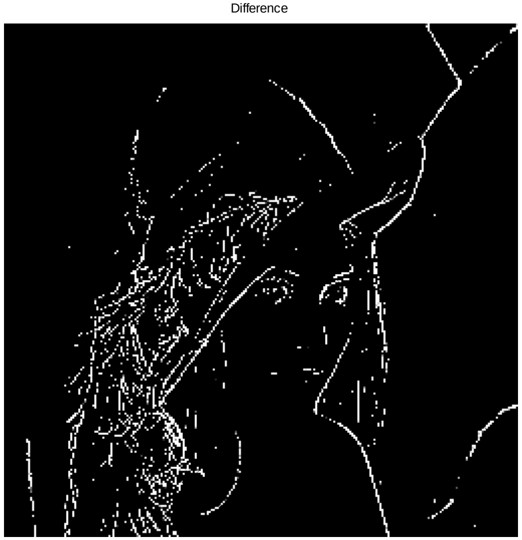
- Perform 2D-DWT to divide image into LL, HL, LH and HH sub-bands.
- Select coefficients from the LL, HL, LH and HH sub-bands that surpass a particular threshold T1
- Embed watermarking data via additive modification

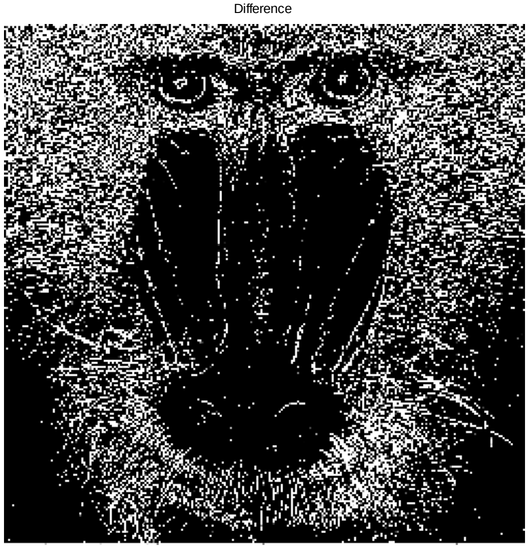
  $t'_i = t_i + \alpha|t_i|x_i$

  x_i = watermark
  α = weighting constant
- Perform 2D-IDWT to create "watermarked image"

Watermark Embedding Method (Cont)

- Modifications to edge data create the least visually perceptible changes
- If using a hard threshold to select coefficients, the number of affected coefficients can vary greatly
- Images with a greater number of edges will hold more watermarking data





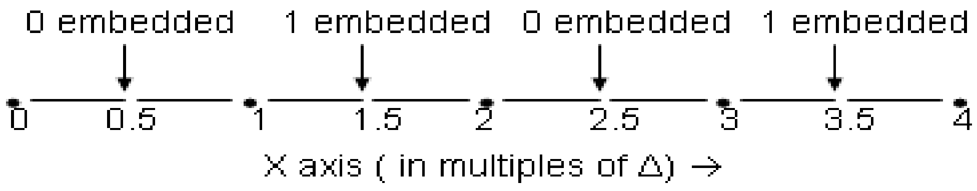
Watermark Detection Scheme

- Method
  - Perform 2D-DWT to divide image into LL, HL, LH and HH sub-bands.
  - Select coefficients from each sub-band that surpass a threshold T2>T1.
  - Compute the correlation z, between the coefficients of the received image (t_i^*) > T2 and a particular watermark (y_i).

$$z = \frac{1}{M}\sum_{i=1} y_i t_i^*$$

Quantization Index Modulation

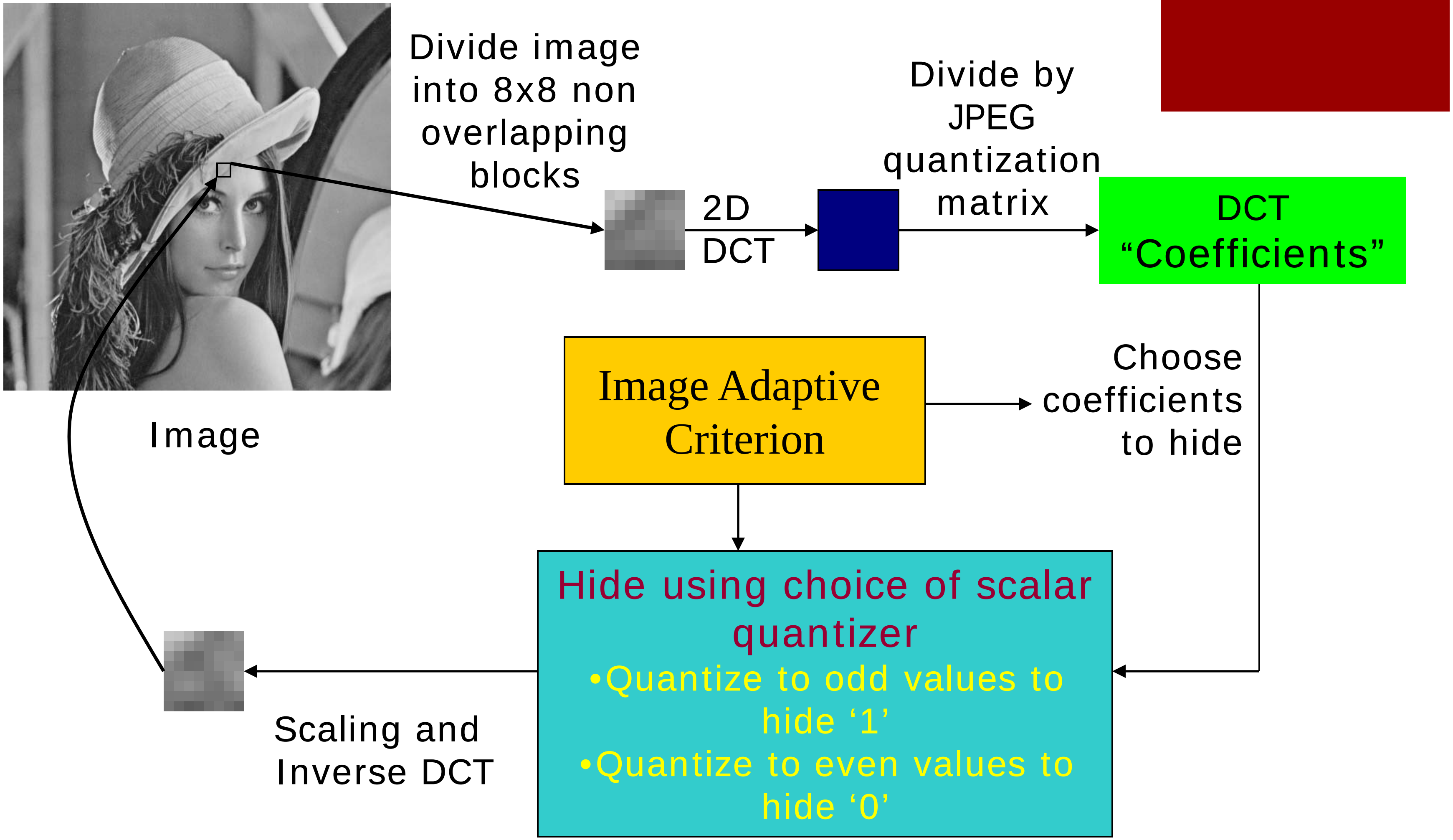
- Message is embedded through choice of quantizer.
- Consider a uniform quantizer of step size ∆ ,Odd reconstruction points represents message ‘1’ & even represents message ‘0’.
- If the value of cover coefficient is ’57.35’ , ∆=4,message bit = ‘1’, then after embedding the message: stego coefficient = 56.



Mapping: a coefficient x is mapped to the nearest x'
where x'= (n+0.5)∆
where n=odd if 1 is embedded
and n=even if 0 is embedded

When |x| < 0.5, it is erased (x is mapped to 0)
When |x| > 0.5, odd-even mapping occurs to hide 0/1

QIM Image Data Hiding



Image

Divide image into 8x8 non overlapping blocks

2D DCT

Divide by JPEG quantization matrix

DCT “Coefficients”

Choose coefficients to hide

Image Adaptive Criterion

Hide using choice of scalar quantizer

- Quantize to odd values to hide ‘1’
- Quantize to even values to hide ‘0’

Scaling and Inverse DCT

Limitations of Digital Watermarking

- Digital watermarking does not prevent copying or distribution.
- Digital watermarking alone is not a complete solution for access/copy control or copyright protection.
- Digital watermarks cannot survive every possible attack.

Challenges in Watermarking Research

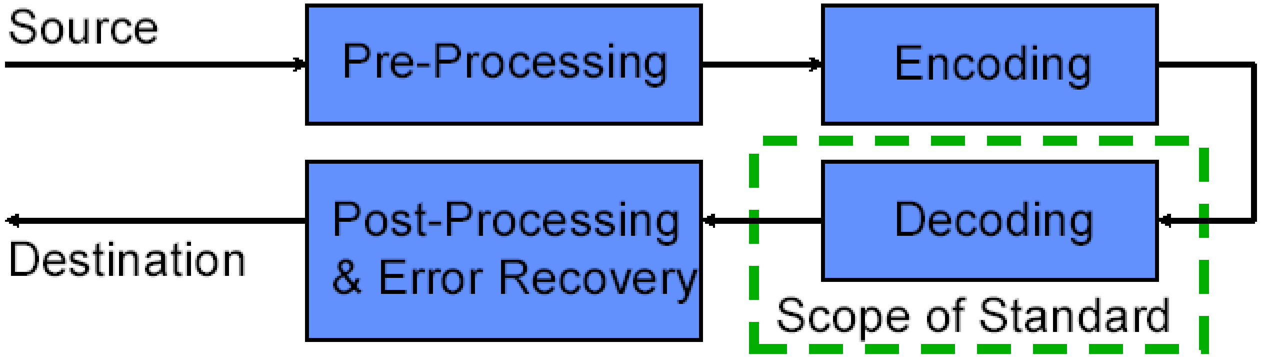
- Lack of protocols, standards and benchmarking.
- Lack of comprehensive mathematical theory.
- Watermark survival for all attacks.
- Relating robustness, capacity, perceptual quality and security.
- Will it be used, and how the legal system adopt it?

Trends in watermarking research

- Color image watermarking, and other multimedia signals.
- 2nd generation watermarking.
- Watermarking of maps graphics and cartoons.
- Information theoretic issues.
- Applications beyond copyright protection.
- Protocols and standardization.

The Scope of Video Coding Standardization

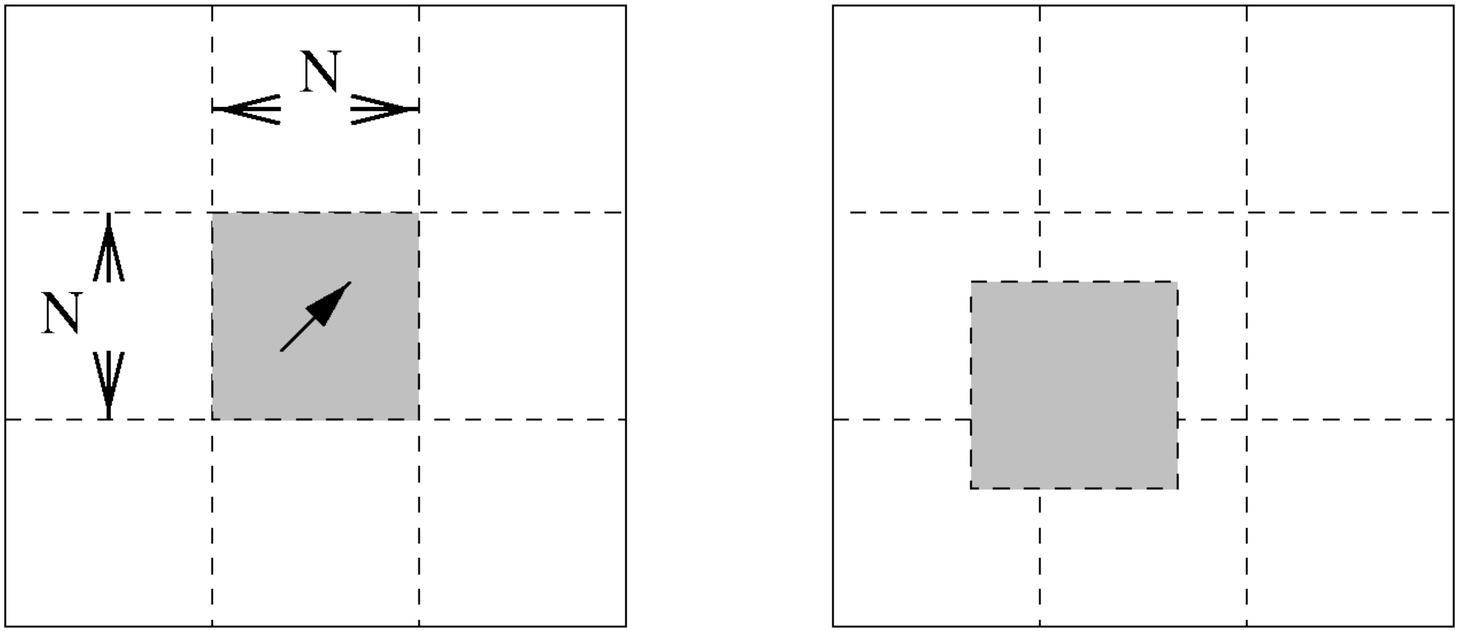
- Only restrictions on the Bitstream, Syntax, and Decoder are standardized:
  - Permits the optimization of encoding
  - Permits complexity reduction for implementability
  - Provides *no* guarantees on quality



Block Motion Estimation

- In a complex nature scene, it is difficult to track all the moving objects.
- Block based motion estimation divides each video frame into fixed-sized non-overlapping blocks, and treat each block as an object to track its motion.
- A block in **current** frame is compared with all possible regions in the **reference** frame, this search will find a best match and produce a **motion vector** representing the displacement between the current block and its best match.
- The reference frame can be previous frame or future frame or both.

Block Motion Estimation



current frame
frame i

previous frame
frame i-1

(a)

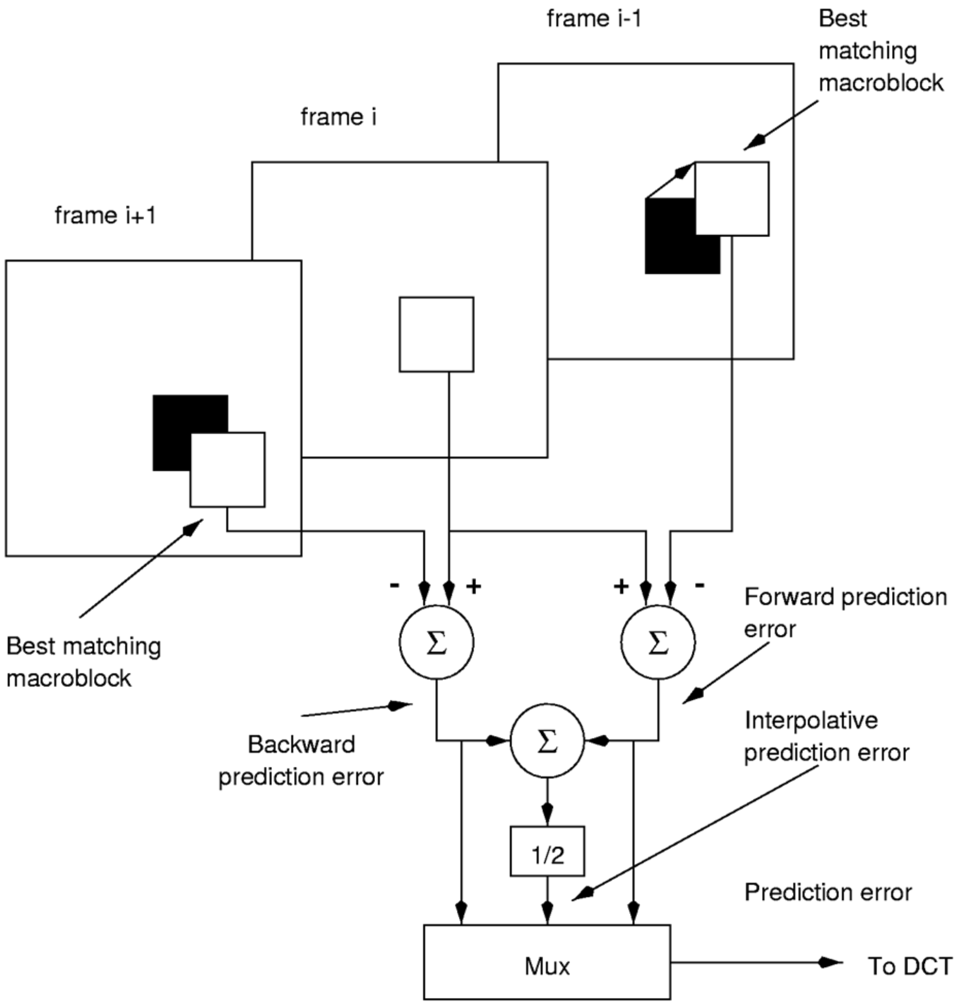
(b)

Motion Compensated Prediction

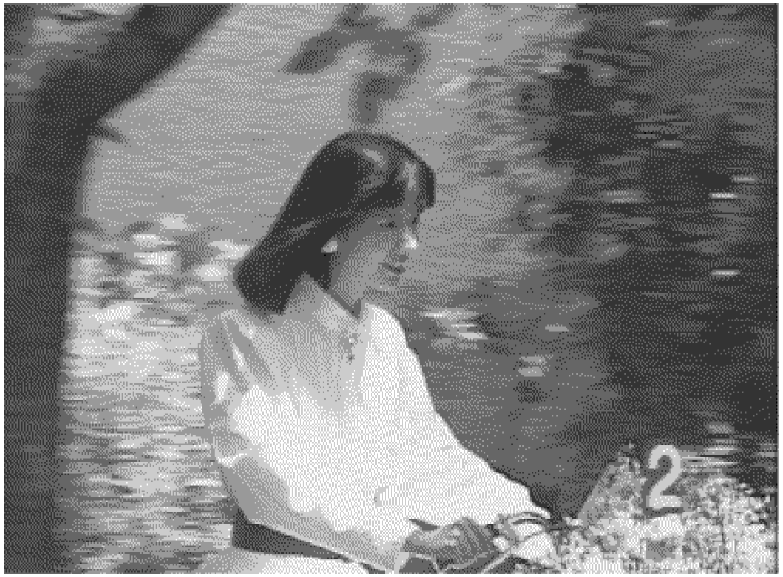
- Motion compensated prediction is a special prediction method based on the knowledge of motion in a video sequence. It can be used in a prediction coding structure.

- The procedure:
  - Block-based motion estimation, generating a motion vector for each block.
  - Moving the best match of each block to its current position to form a prediction block.
  - Deferential coding of each frame based on the prediction frame formed by all the prediction blocks.
  - Sending all the motion vectors to the decoder to generate the same prediction frame for decoding.

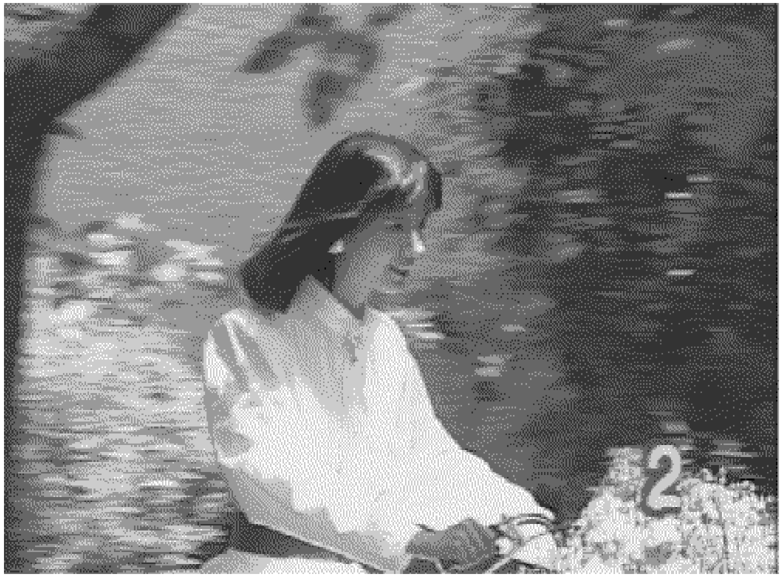
Bi-directional MC Prediction



MC-Prediction: Example

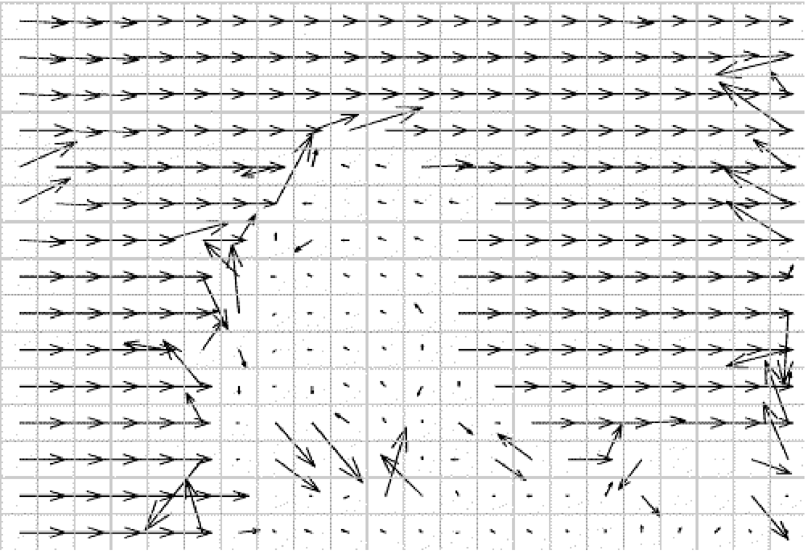


Previous frame

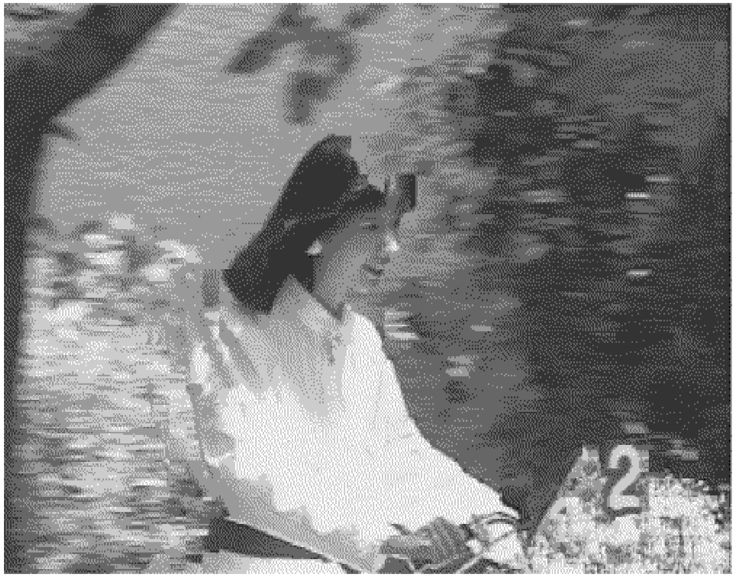


Current frame

MC-Prediction: Example

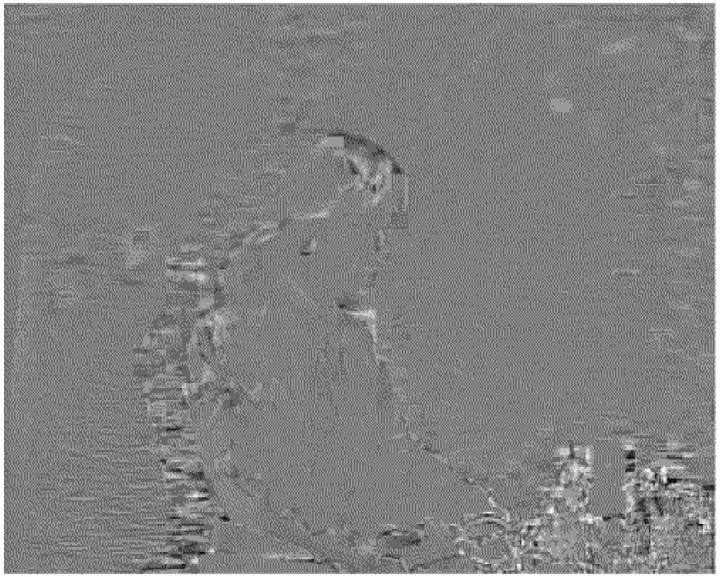


Motion vector field

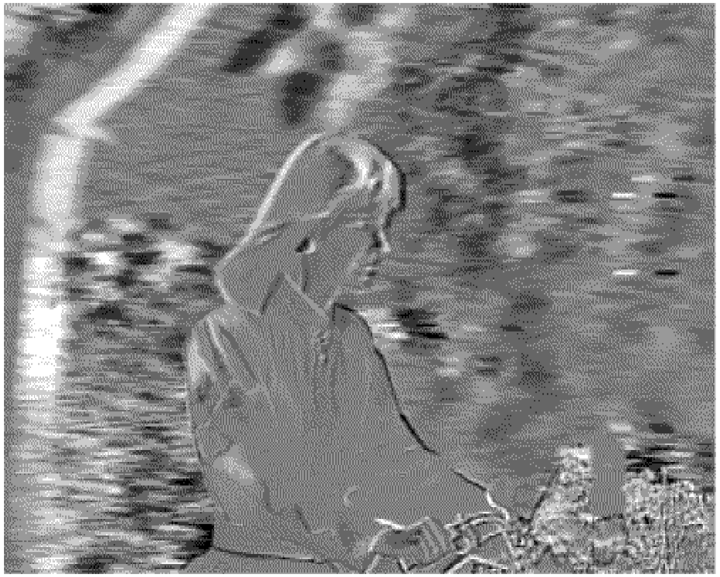


Prediction of current frame

MC-Prediction: Example

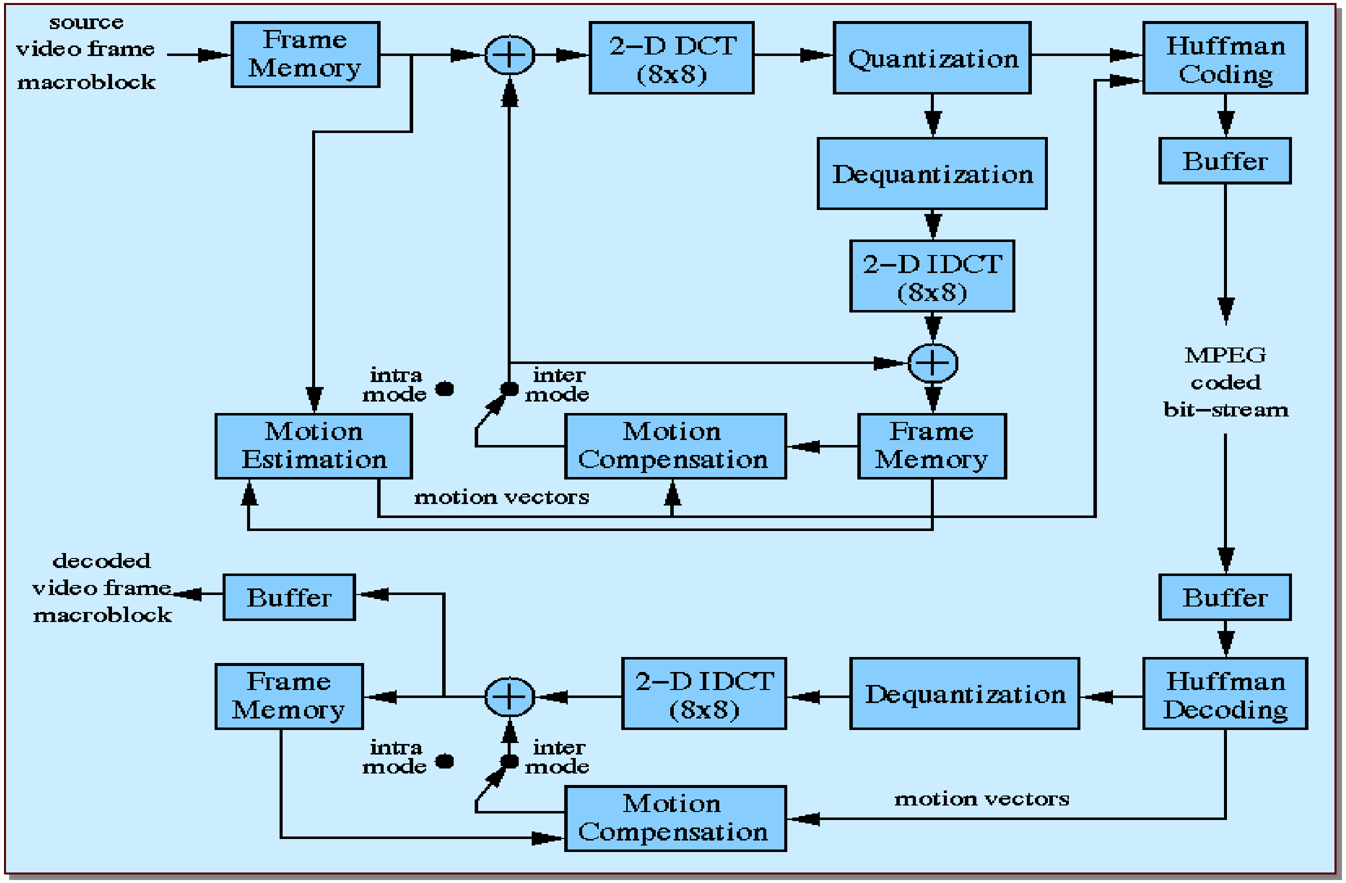


Residual frame with MC



Residual frame with direct difference

Motion Pictures Experts Group (MPEG)

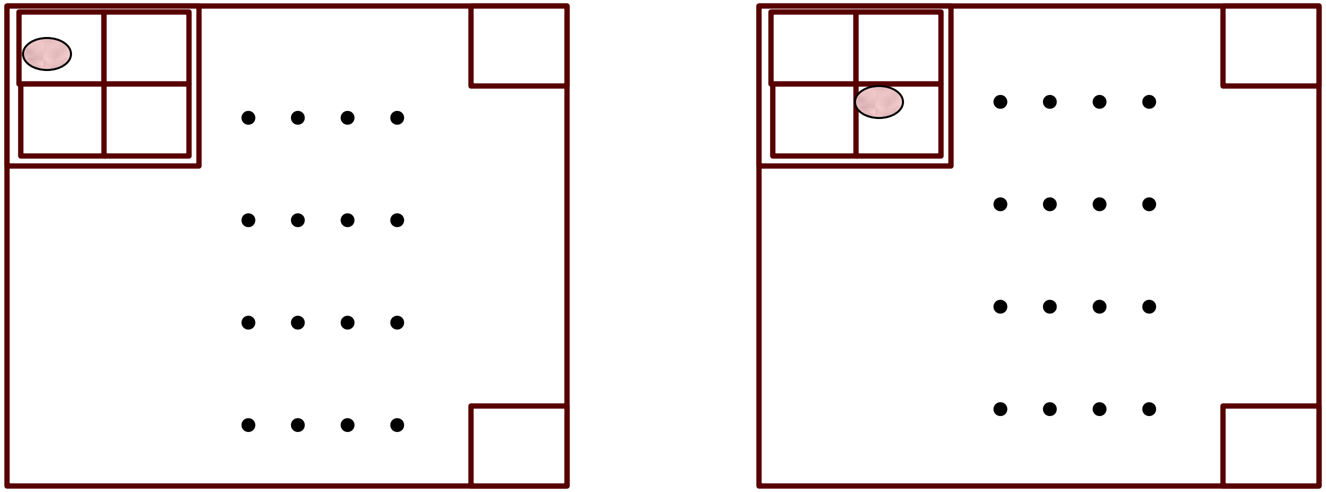


MPEG - I, D frames

- I – frames: self-contained – no other images required for decoding
  - Macroblocks Y (16 x 16), C (8 x 8)
    - 1 Macroblock = 4 (8x8) Y blocks
      + 1 (8x8) C_b block + 1 (8x) C_r block
  - Use JPEG for compresssion
- D – frames: self-contained
  - Save low frequency only
  - Use for fast forward or fast rewind (VCR control simulation)
  - Stored in addition to the regular data stream

MPEG - P frames

- P – frames: require previous I frames.
  - Motion estimation – similar macroblocks in previous I frame
    - Encode motion vector
  - Compute difference block. Use JPEG on difference block.



MPEG - B frames

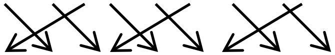
- B – frames: differences based on predictions of previous and next frame
  - 2 motion vectors encoded with respect to previous and next I or P frames
  - Delays the encoding and decoding process
  - Real time and error control considerations, limit the number of B frames used in a sequence

MPEG frame sequences

- I I I I I I I I · · · · · · best reproduction, poor compression
- Frames between I frames (*II*#)
- Prediction span is frames between P and preceding I or P frame (*IP*#)
- I P P I P P I · · · · · *II*# = 3; *IP*# = 1;
- I B B P B B I · · · · · · good reproduction, good compression *II*# = 6; *IP*# = 3;
- Typical values of *II*# 3 to 12; *IP*# 1 to 3;

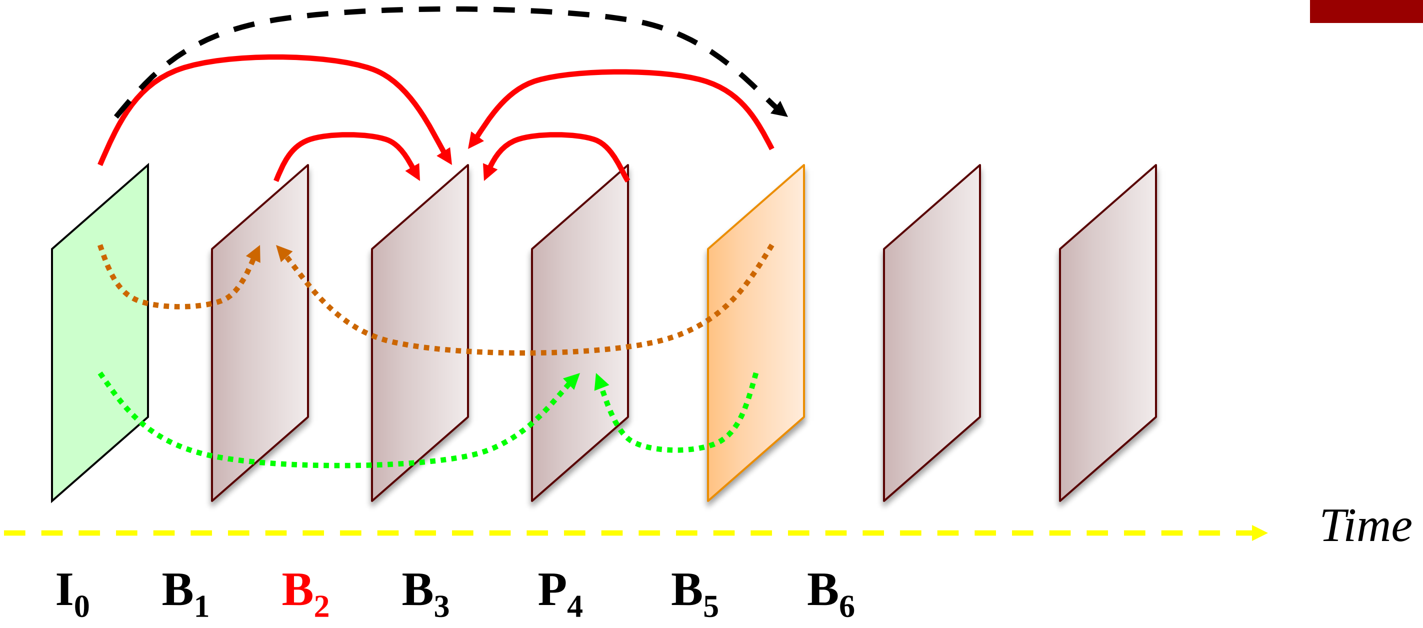
Frame sequence

- Display (capture) order IBBPBBPBBI · · · ·



- Transmission order IPBBPBBIBB · · · ·
- display (capture) order
  - 21 22 23 24 25 26 27 28 29 30
- transmission order
  - 21 24 22 23 27 25 26 30 28 29

B-frame Prediction Weighting



- Playback order: I_0 B_1 B_2 B_3 P_4 B_5 B_6 ………
- Bitstream order: I_0 P_4 B_1 B_3 B_2 P_8 B_5 ………

MPEG-4 Properties

- MPEG-4 defines a set of coding tools and a syntactic description for audio-visual objects (AVOs).
- Each of the AVO can be coded independently.
- Coding of object shapes is essential.
- Object-based motion compensation.
- Panoramic still background is coded using sprite coding.
- Still texture image is coded using wavelet coding.
- Supporting SNR-, spatial- and object-scalability.
- Server-side interaction: content-based manipulations.
- Client-side interaction: AVO modifications.

Video Objects

- Entity that a user is allowed to access and manipulate.
- Area of Video Scene occupying space and time.
- VOP: Instance of VO at a particular point.
- Traditional:

  Each VOP as single frame of video

  VO: sequence of frames.

DVD-Video

- DVD-Video storage capacity 17 Gbyte if two layers on both sides of the disk are utilized.
- Specifications and features
  - 133 minutes of high quality MPEG-2 encoded video with multi-channel Dolby Surround AC-3 audio can be stored on one layer on one side (around 4 Gbyte).
  - Support widescreen, letter box and pan & scan video formats. (4:3 and 16:9 aspect ratios).
  - Up to 8 tracks of digital audio for multiple languages, each with as many as 8 channels.
  - Up to 32 subtitle/karaoke tracks.

DVD-Video

- Specifications and features
  - Menus and program chains for user interactivity
  - Up to 9 camera angles to give the user more choice
  - Digital and analogue copy protection
  - Parental control for protection of children
  - Special effects playback: freeze, step, slow, fast, and scan (no reverse play or reverse step).
  - Random play and repeat play.
  - Programmability (playback of selected sections in a desired sequence).

DVD-Video

- Content Scrambling System (CSS)
  - CSS is a data encryption and authentication scheme intended to prevent copying video files directly from DVD-Video discs
  - The CSS decryption algorithm exchanges keys with the drive unit to generate an encryption key that is then used to obfuscate the exchange of disc keys and title keys that are needed to decrypt data from the disc.
  - DVD players have CSS circuitry that decrypts the data before it's decoded and displayed.
  - In October 1999, the CSS algorithm was cracked and posted on the Internet, triggering endless controversies and legal battles
- DVD regional codes:
  - Motion picture studios want to control the home release of movies in different countries because theater releases aren't simultaneous

Video Watermarking

- Issues on identical watermarks for each frame
  - Problems in maintaining statistical invisibility.
- Issues on independent watermarks for each frame
  - Problems in easy removal of watermarks.
- Robustness:
  - Must survive frame averaging, frame dropping, frame swapping, cropping, temporal rescaling.
  - Must be able to discern imposter watermarks (deadlock). Problems in use of the original video sequence. Problems when no video sequence is needed.

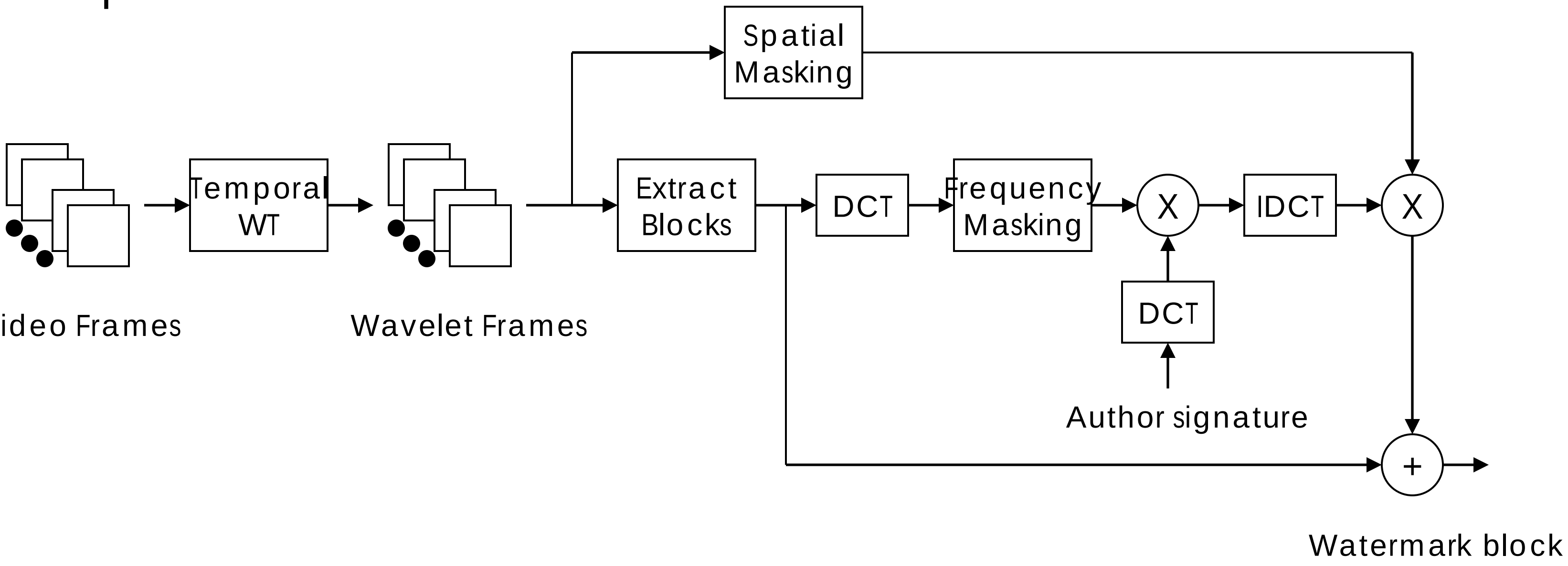
Video Watermarking: A Example (Embedding)

Temporal Wavelet Transform yields:

1) Low-pass frames (Static, non-moving component)

2) High-pass frames (Dynamic, moving component)

Frequency and Spatial Masking are tuned to human visual perception.



Video Watermarking: A Method (Detection)

Detection of Watermark

- With knowledge of location in video sequence

  X = input, R = received coeffs, F = original coeffs, N = noise, W = watermark

  H_0: $X_k = R_k - F_k = N_k$ (No watermark)

  H_1: $X_k = R_k - F_k = W_k + N_k$ (Watermark)

- Without knowledge of location in video sequence (just one frame present)

  Only look at the low-pass frames (static, non-moving component)

- Decision thresholds are determined by a scalar similarity

$$S = \frac{X_k \cdot W_k}{W_k \cdot W_k}$$