CSC7066 / CSC4066
Media Security

Lecture Five

MODELS of WATERMARKING - II

GEOMETRIC MODELS - DEFINITIONS

- Conceptualise the Watermarking Methods
- MEDIA SPACE
  - High dimensional space
  - Each points corresponds to one Work
- MARKING SPACE
  - Projection or distortions of Media Space
  - Analysing more complicated algorithms
- A watermarking system can be viewed in terms of various regions and probability distributions in media or marking spaces

Regions and Distributions

- Distributions of unwatermarked Works
  - How likely each work is
- Region of acceptable fidelity
  - For a given cover Work a region containing all Works which are perceptually similar to a given cover Work
- Detection region
  - Describes the behaviour of the detection algorithms
- Embedding Distortion or Embedding Region
  - Describes the effects of an embedding algorithm
- Distortion Distribution
  - How Works are likely to be distorted during normal usage

Media Space

- WORKS
  - Points in an N-dimensional Media Space
  - N
    - » Graylevel images : Number of pixels
    - » Colour images: 3 x Number of pixels
    - » Audio : number of samples in a segment
    - » Video : Number of frames in a segment x Number of pixels per frame
  - Digital signals
    - » Samples are quantised and bounded : 8-bit (0 - 255)

Distribution of Unwatermarked Works

- Different works have different likelihoods of entering into a watermark embedder or a detector
  - In audio, watermarks are more likely to be embedded into music than into pure static
  - In video, watermarks are more likely to be embedded in images of scenes than in video “snow”

- For a watermarking system, it is important to model the a priori distribution of content we expect the system to process
  - Gaussian distribution
    - General acceptance
  - Laplacian or generalized Gaussian distribution
    - More accurate
  - Result of random, parametric processes
    - More complex, not covered in the textbook

- The Important points
  - The distribution of unwatermarked content is application dependent
  - Accuracy of performance estimation relies on correct choices of distributions of Works

Region of Acceptable Fidelity

- A region in which every vector corresponds to an image that is indistinguishable from c_o
  - It is extremely difficult to determine the exact RAF
  - Depending on HVS about which we know little

- Use some measure of perceptual distance
  - Not necessary a true metric based on HVS

- Mean Square Error (MSE)

$$D_{mse}(c_1, c_2) = \frac{1}{N} \sum_{i=1}^{N} \left(c_1[i] - c_2[i] \right)^2$$

- To define RAF we set a threshold this measure, τ_{MSE} : If $D_{MSE}(c_1,c_2) < \tau_{MSE}$ c_2 is inside in RAF.

- The region of acceptable fidelity becomes an N-dimensional ball of radius $\sqrt{N\tau_{MSE}}$.

- Just Noticeable Difference (JND)
  - More sophisticated

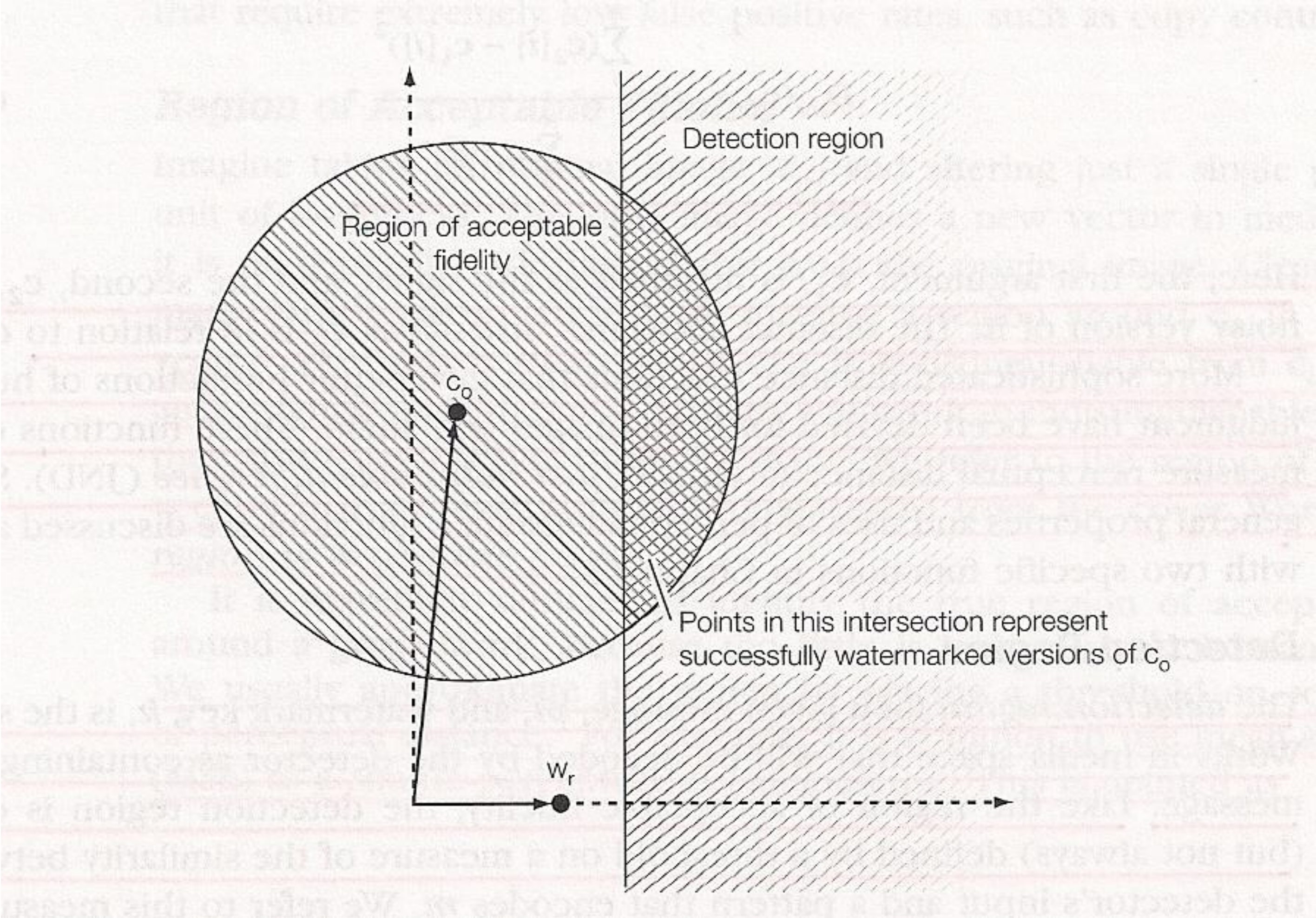
Detection Region

- For a given image and a watermark key, it is the set of Works in Media Space that will be decoded by the detector as containing that message.
- Detection Measure
  - We need similarity measure between the detector's input and a pattern that representing the watermark.
    - » A threshold is also needed to determine the region.
- LC for D_LC algorithm
  - What is the shape of the region?
  - This equals to the product of their Euclidean lengths and cosine of the angle between them.
  - Length of $\mathbf{w_r}$ is constant

$$z_{lc}(c_r, w_r) = \frac{1}{N}(c_r \cdot w_r) = \|c_r\| \|w_r\| \mathrm{Cos}\theta = k\|c_r\| \mathrm{Cos}\theta$$

  - Therefore this measure is to find the orthogonal projection of $\mathbf{c_r}$ on the $\mathbf{w_r}$.

Regions in the Media Space



Embedding Distribution

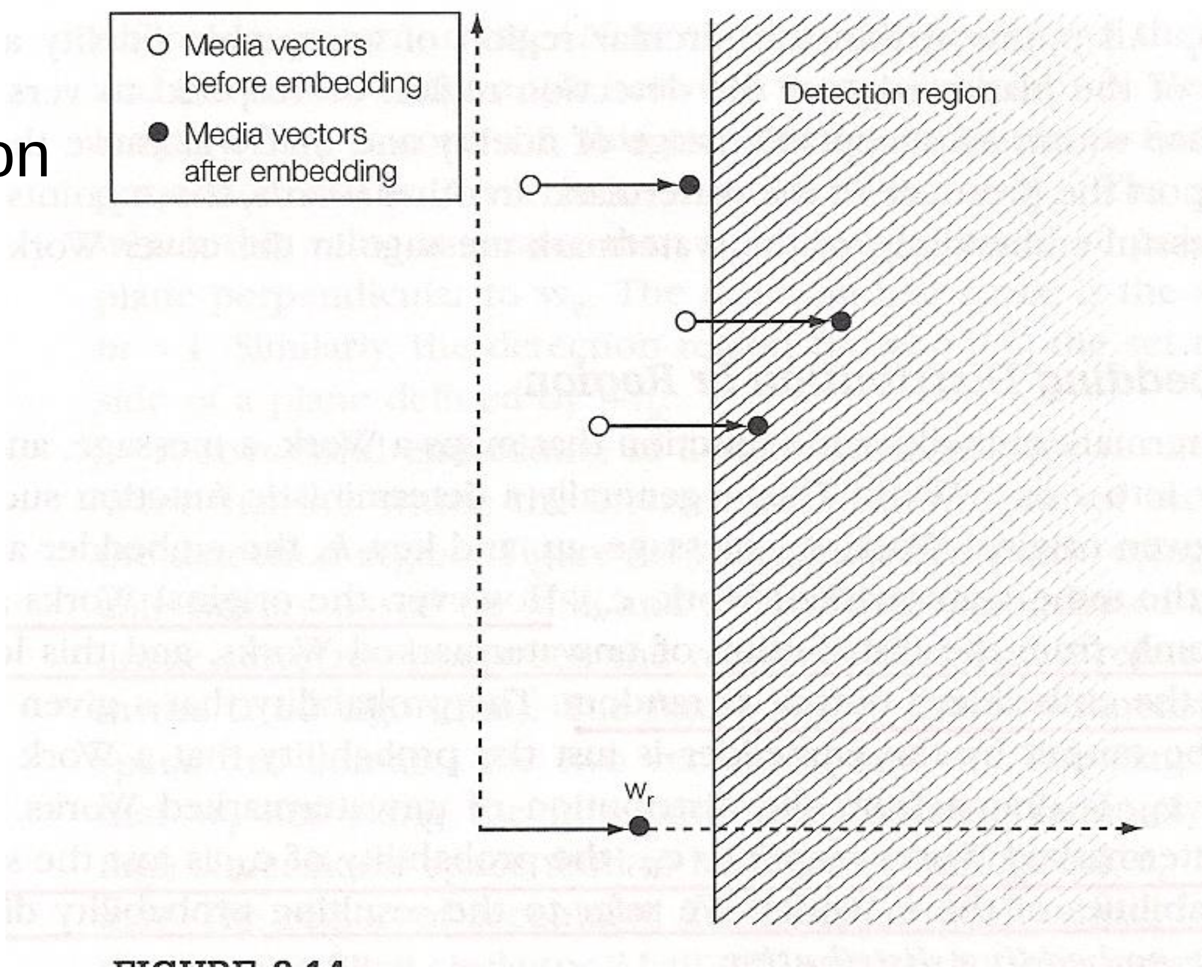
- Embedder
  - Maps a Work and a message into a new Work
  - Generally deterministic function

- Embedding Distribution
  - Original Works are drawn randomly from the distribution of unwatermarked Works.
    - » Embedder's output is random.
  - Probability of $\mathbf{c}_w$ = Probability of $\mathbf{c}_o$

Embedding Distribution

- Some embedding algorithms define an embedding distribution in which every point has a nonzero probability.

- E_BLIND



Improving E_BLIND

- Using Side Information
- E_FIXED_LC

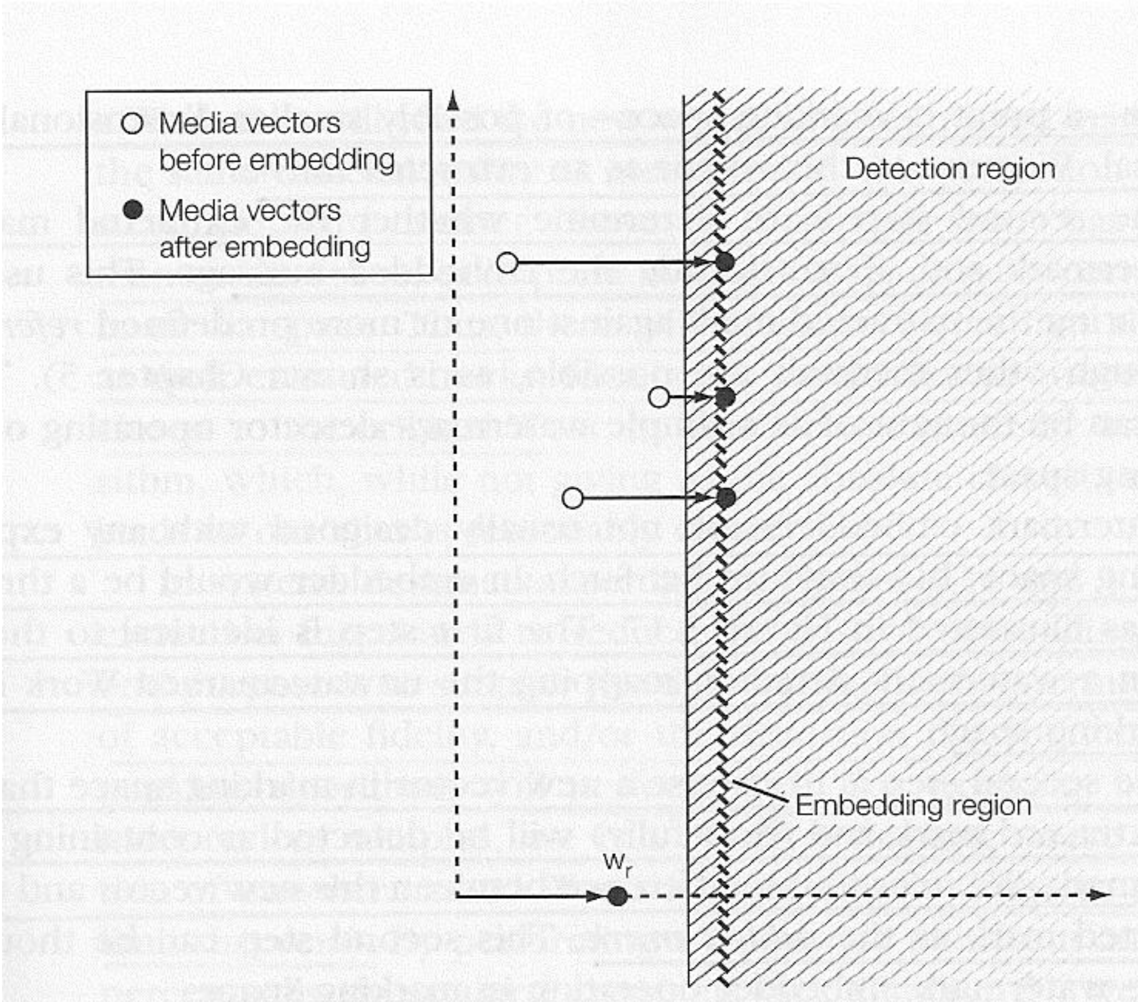


FIGURE 3.15

Distortion Region

- Probability of obtaining a given distorted Work $\mathbf{c}_{wn}$ depending to given undistorted watermarked Work $\mathbf{c}_w$.

$$P(c_{wn}|c_n)$$

  - Same type of distribution used in modelling transmission channel.
- Assumption: Additive Gaussian Channel
  - Simplifies the analysis; not the exact case
  - Some attacks depend on the content
    - » Compression, cropping, rotation

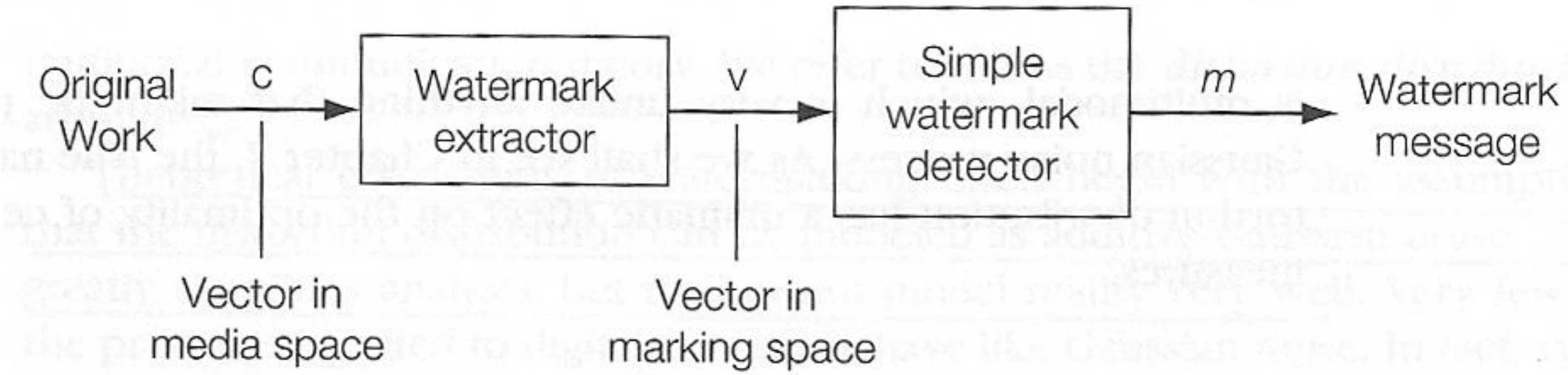
Marking Spaces

- For E_BLIND/D_LC and E_FIXED_LC/D_LC methods identifying the embedding and the detection regions are not very difficult.

- It is not always true for more complex methods.
  - It is useful to view part of the system as performing a projection or distortion of Media Space into a Marking Space.

Marking Space - Detector

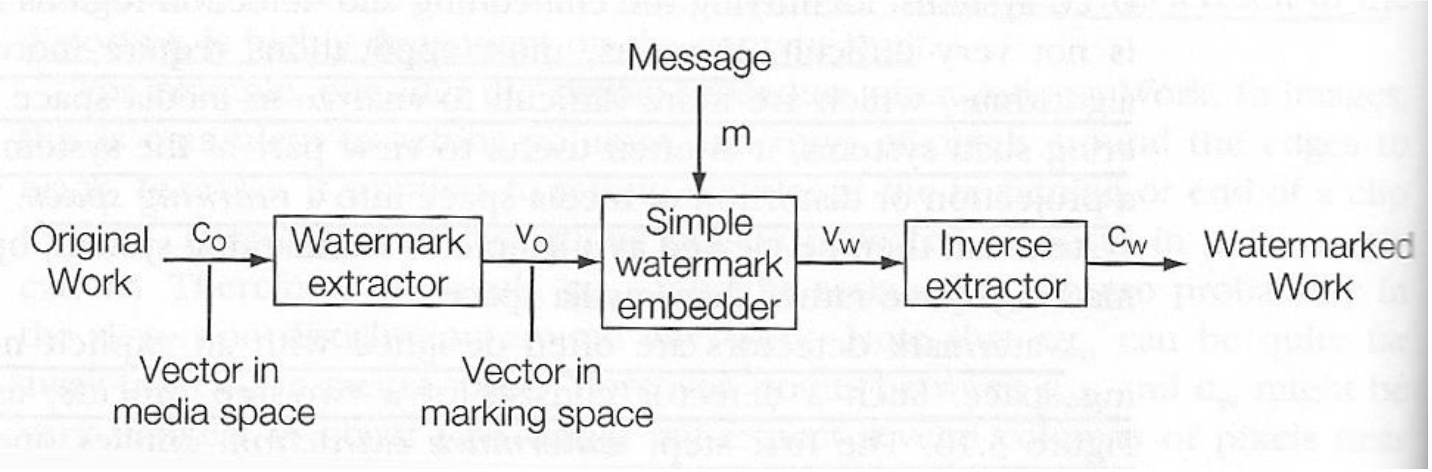
- Designed with an explicit notion of watermarking space
- Two step process



- First Step
  - » Preprocessing of the content ; frequency transforms, filtering, averaging,…
  - » Result is a vector : Extracted watermark
    - Smaller dimensionality than the original
- Second Step
  - » Determine whether the extracted mark contains a watermark or not

Marking Space - Embedder

- Not usually designed with any explicit use of marking space
  - But they can be
- Three step process



  - **First Step** : mapping the unwatermarked work into a point in marking space
  - **Second Step** : choose a new vector in marking space that is close to the extracted mark and will be detected
  - **Third Step** : Invert the process to go back to the media Space to obtain the watermarked work.

Block-based Blind Embedding & Correlation Coefficient Detection

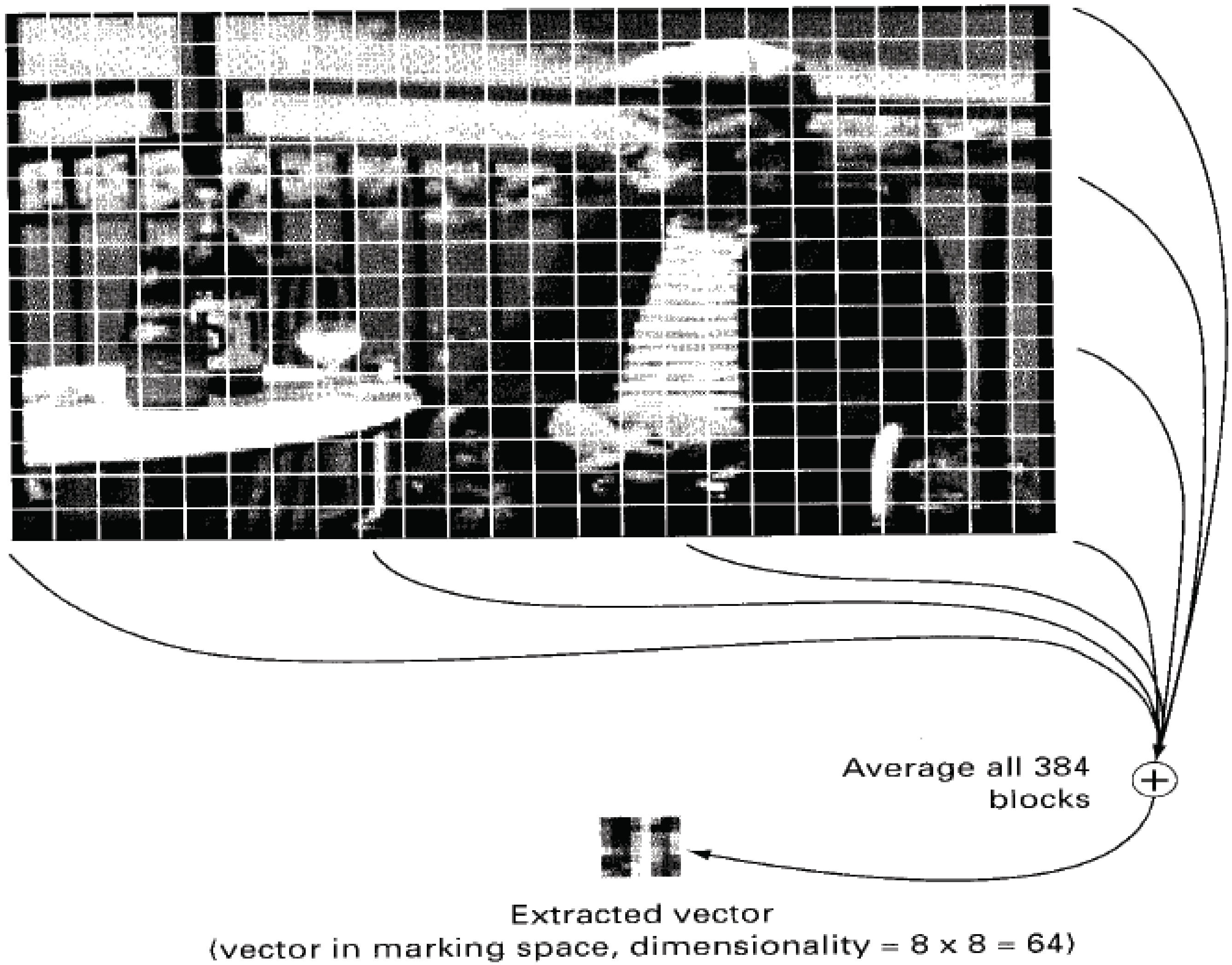
- Extract Watermarks by averaging 8x8 blocks of the image
  - 64-dimensional marking space
- Embedding in marking space by simple, blind embedding
  - Resulting changes are projected back to the full size
  - Single bit embedding
- ***Correlation coefficient*** is used in the detector.
  - Normalised version of the Linear Correlation
  - Watermark/ No watermark decision

Detector - I

- Two steps
  - Extract a mark **v**, from the received Work, **c**.
  - Use a simple detection algorithm to detect a watermark in the extracted mark, v.

- STEP 1 :
  - Divide the image into 8x8 blocks
  - Average all blocks into one array of 64 values.

Detector -II



Average all 384
blocks
Extracted vector
(vector in marking space, dimensionality = 8 x 8 = 64)

Detector - III

- Step 2:
  - **Correlation Coefficients**
    - » What is about Linear Correlation?
    - » Correlation coefficient is more robust against to certain distortions.
  - Differ from Linear Correlations
    - » Subtract the means of the two vectors before correlating them.
      - The detection value is unaffected if a constant is added to all elements of either of the two vectors.
    - » Normalise the linear correlation by the magnitudes of the two vectors.
      - The detection value is unaffected if all elements of either of vector are multiplied by a constant.
  - Correlation Coefficient is robust to changes in image brightness and contrast.

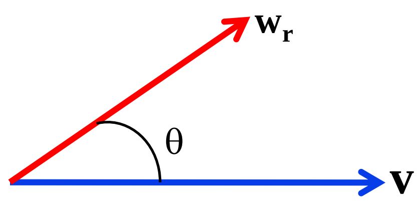
Detector - IV

- Correlation Coefficient

$$z_{cc}(\mathrm{v},\mathrm{w}_r)=\frac{z_{lc}(\tilde{\mathrm{v}},\tilde{\mathrm{w}}_r)}{\sqrt{z_{lc}(\tilde{\mathrm{v}},\tilde{\mathrm{v}})\times z_{lc}(\tilde{\mathrm{w}}_r,\tilde{\mathrm{w}}_r)}}\quad ;\ \tilde{\mathrm{v}}=\mathrm{v}-\mathrm{v}_{mean}\qquad \tilde{\mathrm{w}}_r=\mathrm{w}_r-(\mathrm{w}_r)_{mean}$$

$$z_{lc}(\tilde{\mathrm{v}},\tilde{\mathrm{w}}_r)=\sum_{i=1}^{8}\sum_{j=1}^{8}\tilde{\mathrm{v}}[i,j]\times\tilde{\mathrm{w}}_r[i,j]$$

- Inner product of **v** and $\mathbf{w_r}$ after normalisation
- Cosine of angle between two vectors



$$z_{cc}(\mathrm{v},\mathrm{w}_r)=\mathrm{Cos}\theta$$

- Bounded : $-1\le z_{cc}(\mathrm{v},\mathrm{w}_r)\le 1$

Detector - V

- D_BLK_CC

$$m_n = \begin{cases} 1 & \text{if} \quad z_{cc}(v, w_r) > \tau_{cc} \\ \text{no watermark} & \text{if} \quad -\tau_{cc} \le z_{cc}(v, w_r) \le \tau_{cc} \\ 0 & \text{if} \quad z_{cc}(v, w_r) < -\tau_{cc} \end{cases}$$

  - τ_{cc} is a constant threshold.

Embedding

- Blind embedding in Marking space
  - Similar to E_BLIND

- Added Watermark is the same

$$w_m = \begin{cases} w_r & \text{If } m = 1 \\ -w_r & \text{If } m = 0 \end{cases} \qquad w_a = \alpha \times w_m$$

- Adding in Marking space

$$v_w = v_o + w_a$$

- Projections

$$c_o \xrightarrow{\text{to marking space}} v_o \qquad v_w \xrightarrow{\text{to media space}} c_w$$

  - $\mathbf{c_o}$ and $\mathbf{c_w}$ should be perceptually similar.

Embedding - II

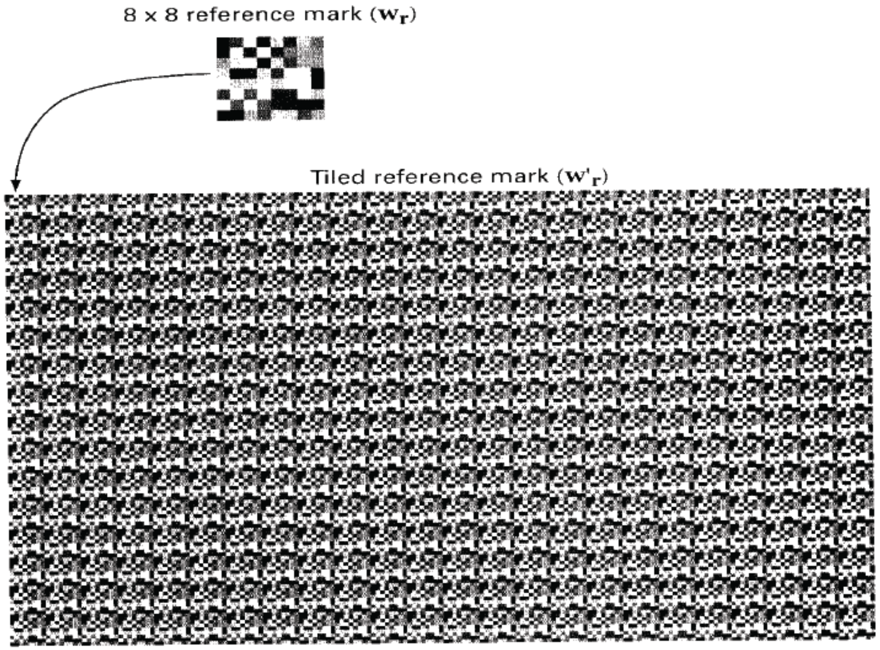
- Problem
  - The extraction is many-to-one
- Solution
  - Distribute each element of the desired change in the extracted mark uniformly to all contributing pixels.

$$c_w[x,y] = c_o[x,y] + \left(v_w[x \bmod 8, y \bmod 8] - v_o[x \bmod 8, y \bmod 8]\right)$$

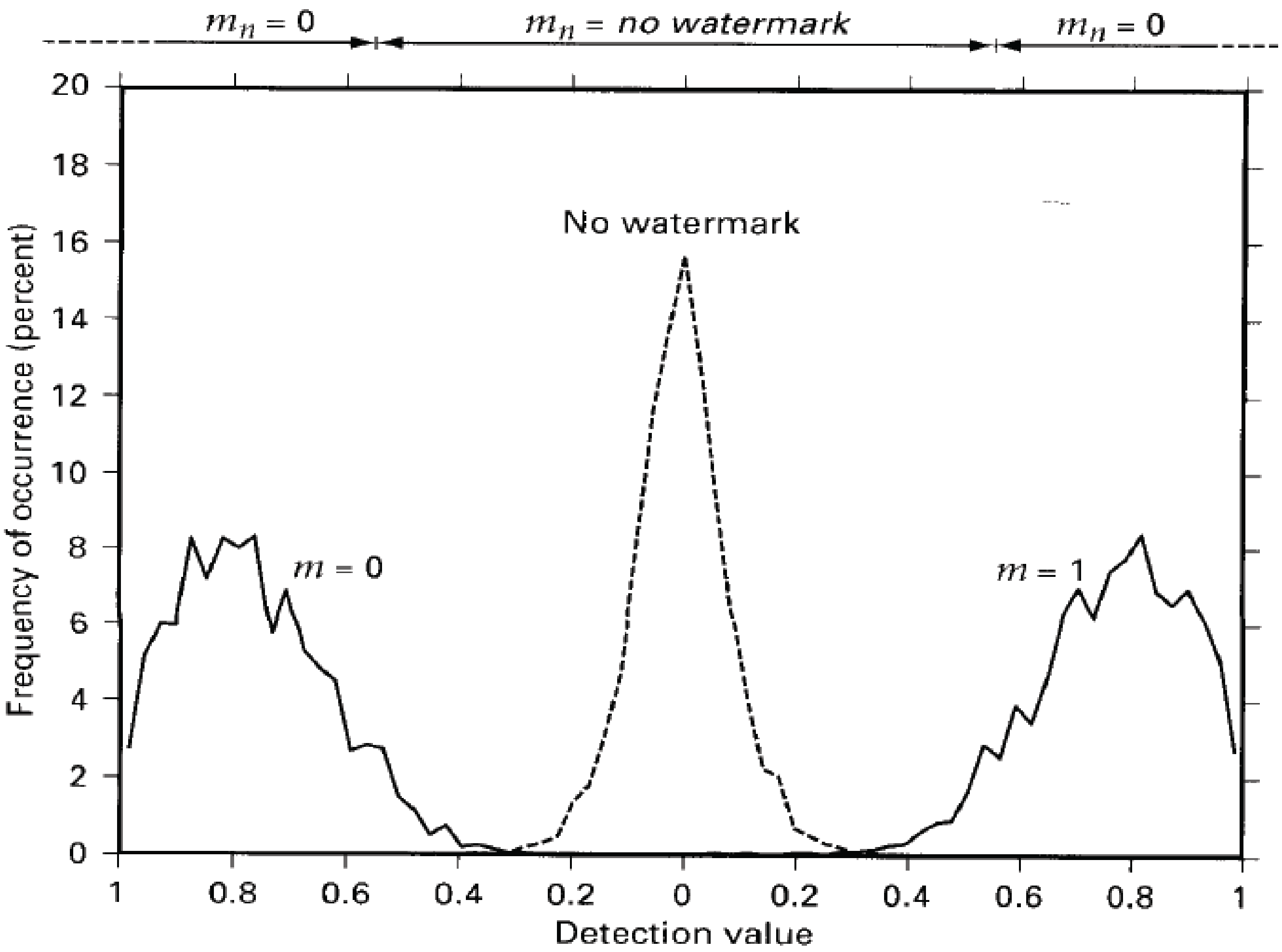
  - Remember vector in the marking space has sizes of 8x8.
  - mod : modulo operator

Embedding - III

- If we use LC rather than CC, performance of this system would be identical to the previous one.
  - Reference pattern consists of a single 8x8 pattern tiled over the full size of the image
    - » $w'_r[x,y] = w_r[x \bmod 8, y \bmod 8]$



Result with CC



Comparison : E_Blind/D_LC vs. E_Blk_Blind/ D_Blk_CC

- Advantages of D_Blk_CC
  - Robust against to certain changes in image brightness and contrast
  - Computationally cheaper that the D_LC

- Disadvantages of D_Blk_CC
  - The number of possible reference marks is much smaller
    - » 256 x 256 $\rightarrow$ 8x8
    - » Lead to poor statistical performance
      - Randomly generated reference marks tend not to work well
      - Must be carefully selected
  - Very sensitive to clipping and rounding
    - » Error diffusion techniques should be used to avoid cumulative error.