Spread Spectrum Watermarking

Information Technologies for IPR Protection
• A Narrow-band signal is transmitted over a much larger bandwidth such that the signal energy presented in any signal frequency is undetectable.
• A watermark is spread over many frequency bins so that the energy in one bin is very small and certainly undetectable.
Because the watermark verification process knows the location and content of the watermark, it is possible to concentrate these weak signals into a single output with high SNR.
• Remark:
To destroy such a watermark would require noise of high amplitude to be added to all frequency bins.
Why it is secure?

- The location of the watermark is not obvious

- Frequency regions should be selected that ensures degradation of the original data following any attack on the watermark
• The energy in the watermark is sufficiently small in any single frequency coefficient.

• It is possible to increase the energy present in particular frequencies by exploiting knowledge of masking phenomena in human auditory and visual systems.
Upon applying a frequency transformation (FFT/DCT) to the data, a perceptual mask is computed that highlights perceptual significant regions in the spectrum that can support the watermark without affecting perceptual fidelity.
Why it is robust? (1/2)

- To be confident of eliminating a watermark, an attack must attack all possible frequency bins with modifications of certain strength → create visible defects in the data
Unintentional signal distortions due to compression or image manipulation, must leave the perceptually significant spectra components intact, otherwise the resulting image will be severely degraded.
• In order to place a length $n$ watermark into an $N \times N$ image, the $N \times N$ DCT of the image is computed and watermark is placed into the $n$ highest magnitude coefficients (which are data dependent) of the transformed image, excluding the DC component (not necessary).
A watermark consists of a sequence of real numbers $\mathcal{X} = x_1, \ldots, x_n$, where each value $x_i$ is chosen independently according to $N(0, 1)$: normal distribution assumption.
• When we insert $\mathcal{X}$ into $V$ to obtain $V'$ we specify a scaling parameter $\alpha$, which determines the extent to which $\mathcal{X}$ alters $V$. 
(1) \( v'_i = v_i + \alpha x_i \)

(2) \( v'_i = v_i (1 + \alpha x_i) = v_i + \alpha x_i v_i \)

(3) \( v'_i = v_i (e^{\alpha x_i}) \) or \( \log v'_i = \log v_i + \alpha x_i \)

- (1) may not be appropriate when the values of \( v_i \) varying widely!!
- Insertion based on (2) or (3) are more robust against ‘different in scales.’
A single $\alpha$ may not be applicable for perturbing all of the values $\nu_i$, since different spectral components may exhibit more or less tolerance to modification.
\[ \nu_i' = \nu_i \left( 1 + \alpha_i x_i \right) \]

Where \( \alpha_i \) can be viewed as a relative measure of how much one must alter \( \nu_i \) to alter the perceptual quality of the host.
• A large $\alpha_i$ means that one can perceptually ‘get away’ with altering $v_i$ by a large factor without degrading the host.

• (2) can be viewed as a special case of the generalized (1) ($v_i' = v_i (1 + \alpha_i x_i)$), for $\alpha_i = \alpha v_i$
A large value is less sensitive to additive alternations than a small value.
Question
How sensitive an image to various additive alternations is highly related to Human perceptual systems.

- Subjective distance measurement: JND
- Objective distance measurement
  \[ \alpha_j \sim \text{median} \left| \nu_j^* - \nu_j \right| \]
  \[ \text{max} \]
• The choice of $n$ indicated the degree to which the watermark is spread out among the relevant components of the image.

• In general, as the number of altered components are increased the extent to which they must be altered decreases — the ability to against attacks ↓
Choice of $n$ (2/2)

- Information theoretic assessment
  - Channel Capacity
  - Rate-Distortion Theory
Evaluating the Similarity : Detection scheme

- $X^*$: the extracted mark
- $Sim(X, X^*) = (X^* \cdot X) / \sqrt{X^* \cdot X^*}$
- $X$ and $X^*$ match, if $Sim(X, X^*) > T$
- Minimize the rate of
  - False Negatives (missed detections)
  - False Positive (false alarms)
Other Issues

• Is there any post processing that can improve the watermark detection performance?

  – Statistical approach
  – Optimization approach: NN, GA
  – With the aid of the knowledge of the host: Communication with side information
  – Game Theory
Attacks (1/2)

- Collision attack
- Image Scaling
- JPEG Compression
- Dithering Distortion
- Cropping attack
- Print, Xerox, and Scan
Attacks (2/2)

- Attack by Watermarking Watermarked Image
  - multiple watermarking

Sometimes, this is a must
Sometimes, this produces ‘Watermark Dilemma’