CSC7066/CSC4066
Media Security

Tutorial One

SPREAD SPECTRUM WATERMARKING for MULTIMEDIA

by Cox et al.

Generic Research Paper

- Introduction
  - Brief explanation of the problem
    - » You are writing to the experts or to the researchers who are familiar with the subject.
  - Key ideas behind the research
  - Previous work if there is no section for it
- Previous Work
  - If there is no mentioned in the introduction
- Presentation of the Research
  - Break down into sections and subsections to make it easy to follow
  - Clear presentation
  - References to the previous works
- Results or Experiments
  - Good explanation of the experiments
    - » Others should repeat the experiments
  - Evaluation and Explanation of the results
- Conclusion
  - Appropriate conclusions from the work should be drawn.
- References

Cox's paper

- What are the key points of the paper?
  - From Introduction
    - Watermark Structure
    - Insertion Strategy
  - Watermark structure
    - Independent Identically Distributed (i.i.d) Gaussian signal ; N(0,1)
      - Collusion attack
      - Quantisation attack
      - Structured to low false positive and false negative detection
  - Insertion Strategy
    - Perceptually most significant **spectral** components
      - Without perceptual degradation
      - Most signal processing techniques leave them intact
    - Spectra means that it is in the transform domain
      - Discrete Cosine Transform

Cox’s paper

- What is the weakness of the paper?
  - Normally the weaknesses will be discovered by other authors later on.
  - Sometimes the authors pointed out the weaknesses as well.
  - **NO PROOF of CONTENT OWNERSHIP !!!**
    - » No countermeasure against watermark insertion.
  - HOWEVER THIS IS A KEY PAPER
    - » **Spread Spectrum Concept**

Spectral (Frequency) Domain Watermarking

- Common Attacks
  - Lossy Compression
    - Eliminates high frequency components. WHY?
      - Human Visual System (HVS) is less sensitive to high frequency components.
  - Geometric distortions
    - Spectral domain spreads the watermark over the whole spatial domain
  - Other attacks
    - ????
- Conclusion
  - **Difficult to find a solution for all type of attacks**

Where is the idea coming from?

- ❑ Spectral or frequency domain → Communication Channel

- ❑ Watermark → Signal to be transmitted

- ❑ Attacks → Noise

- ❑ **SPREAD SPECTRUM COMMUNICATION**
  - Transmit a narrowband signal over a much larger band signal
    - » Signal energy present in any single frequency is undetectable
  - Watermark : narrow band ; image : larger band
    - » Spread watermark over very many frequency bins of image spectra
    - » Small energy ; cannot be detectable

Which spectral bands?

- ❑ Fourier transform (FFT)
- ❑ Discrete Cosine Transform (DCT)
- ❑ Discrete Wavelet Transform (DWT)

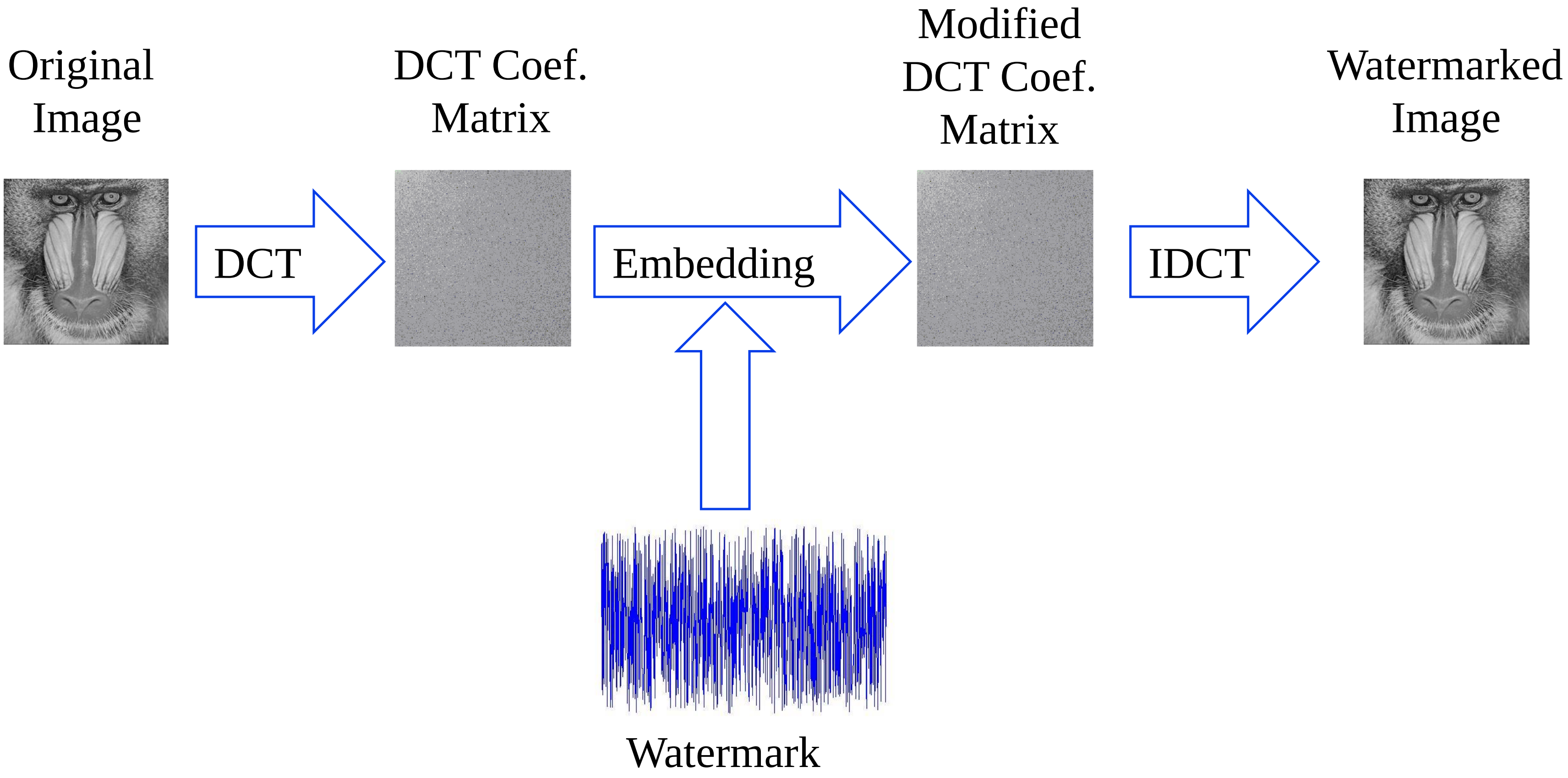
Method

- Watermark
  - Gaussian N(0,1) ; length ***n***

- DCT of whole image → DCT coefficient matrix

- Insert watermark into the ***n*** highest magnitude coefficients of the transform matrix, **excluding the DC component**.
  - WHY?

Method (Cont.)



Original Image

DCT

DCT Coef. Matrix

Embedding

Modified DCT Coef. Matrix

IDCT

Watermarked Image

Watermark

Embedding

- Different embeddings
  - $v_i' = v_i + \alpha x_i$

    $v_i' = v_i(1 + \alpha x_i)$ ←

    $v_i' = v_i(e^{\alpha x_i})$

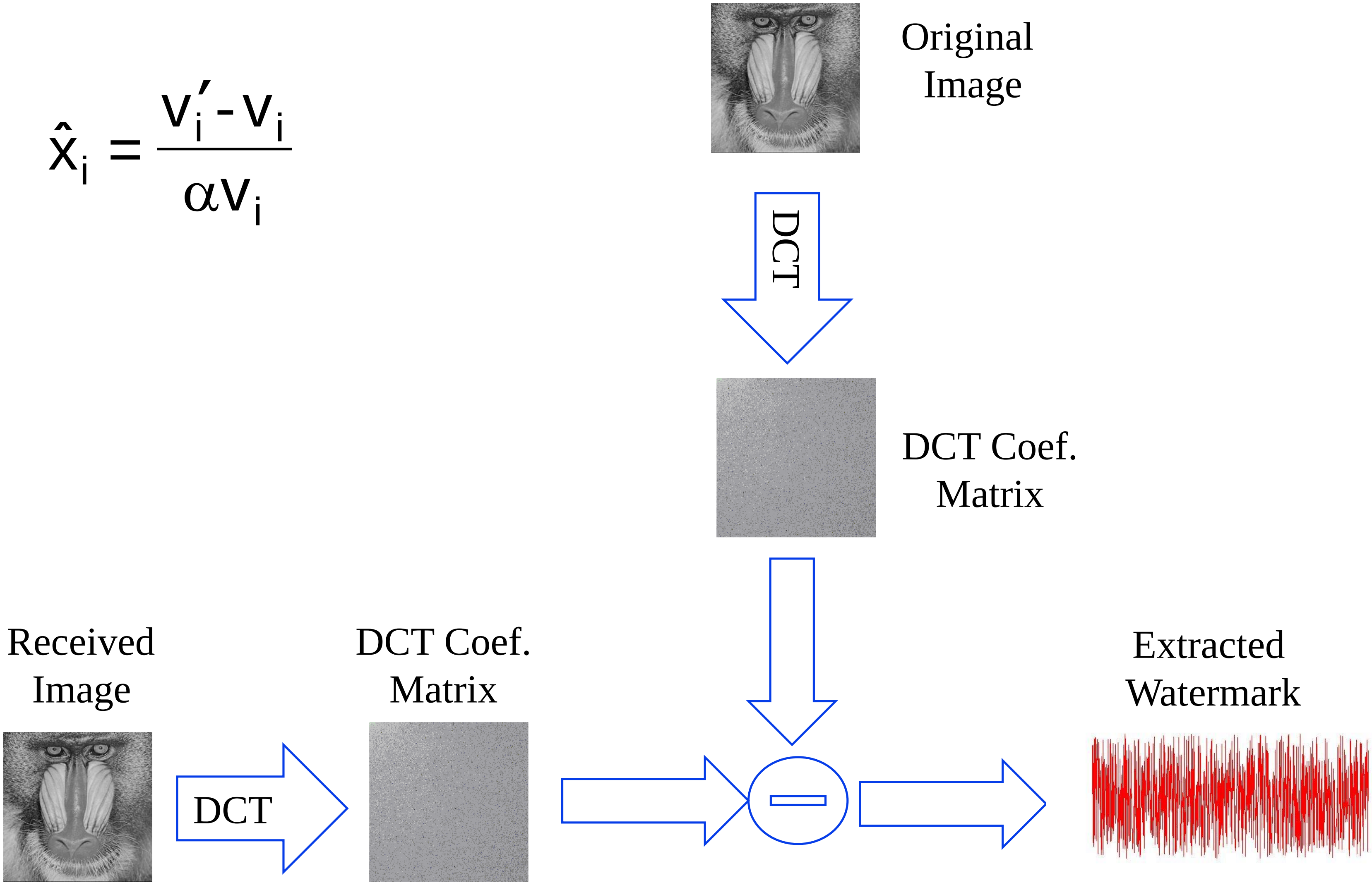
  - What is α ?
    - » Strength parameter: Determines the trade off between robustness and the fidelity

  - What are the properties of these three approaches?

  - Which one was used in the paper?

Extracting Watermark

$$\hat{x}_i = \frac{v_i' - v_i}{\alpha v_i}$$



Original Image

DCT

DCT Coef. Matrix

Received Image

DCT

DCT Coef. Matrix

Extracted Watermark

Similarity of Watermarks

- FACT : Original watermark and the recovered one cannot be the same.
  - WHY?

- Measure: Correlation Coefficient

$$\mathrm{sim}(X, X^*) = \frac{X^* \bullet X}{\sqrt{X^* \bullet X^*}}$$

- sim(X,X*) > Thr $\Rightarrow$ Watermark is there!!!

Similarity of Watermark (cont.)

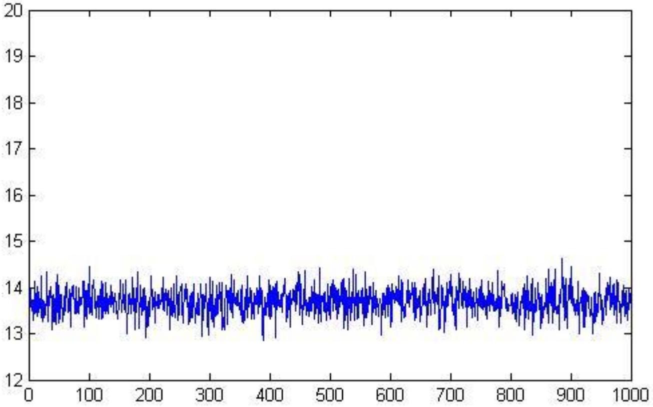
- Determining the threshold is not a trivial problem

- Empirical observation

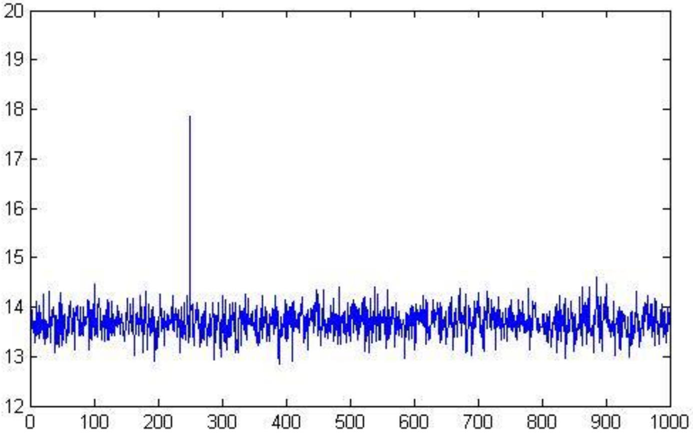
  - Generate a large number of different watermarks
  - Insert the original one into this set
  - Calculate Sim(X,X*) for all watermarks

  - If the original watermark presents in the image, its Sim value should be significantly larger than the other Sim values

Interpretation



NO WATERMARK



WATERMARK is THERE