Watermarking

By Guy-Raphaël Stauffer, David Alexander and Chevalley Gibran

Context on the subject

Watermarking vs. Steganography

- Steganography: transmit a message by hiding it in another message
- Watermarking: embed permanently a watermark in another message
  - Removing the watermark would deteriorate the message
  - Must be resistant to message modification (Robustness)
  - Usages :
    - Owner identification
    - Authentication
    - Labelling
    - Tracing broadcast

Watermarking

Visible vs Invisible

Visible watermarks examples:

Art

https://www.masswatermark.com/wp-content/uploads/2018/03/Canoe.jpg

Television

https://www.rts.ch/2016/05/07/20/17/7705395.image

Banknotes

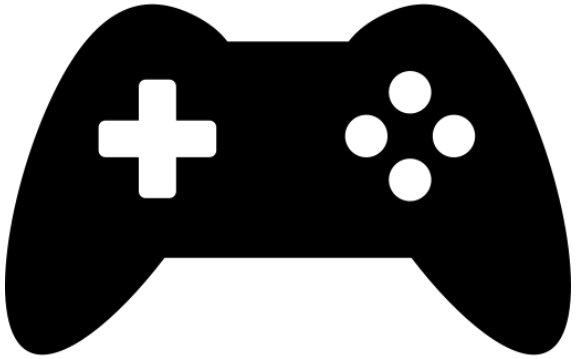
https://www.snb.ch/fr/mmr/reference/nb_20/source/nb_20.fr.pdf

Watermarking

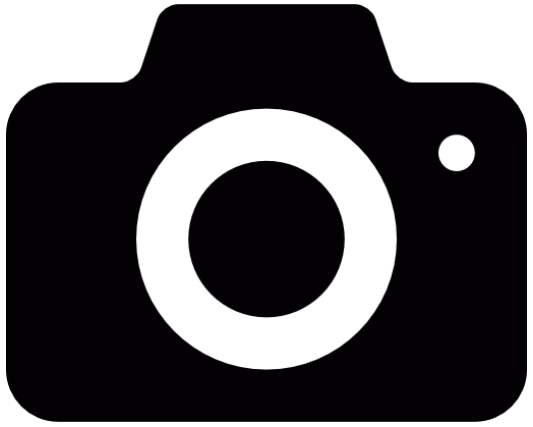
Visible vs Invisible

Invisible watermarks examples:

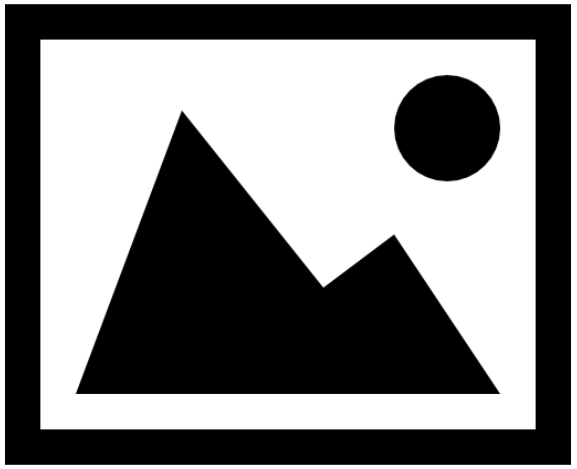
Video games



Camera



Shooting pictures



Time domain methods

Brief overview of different methods

- Hide the watermark in least significant bit the of each pixel
- Use geometric paterns
- Hide information in the difference of intensity of two pixels
- Use blocs of pixels instead of only one to be more resistant against JPEG compression

Keypoints of my method to improve robustness

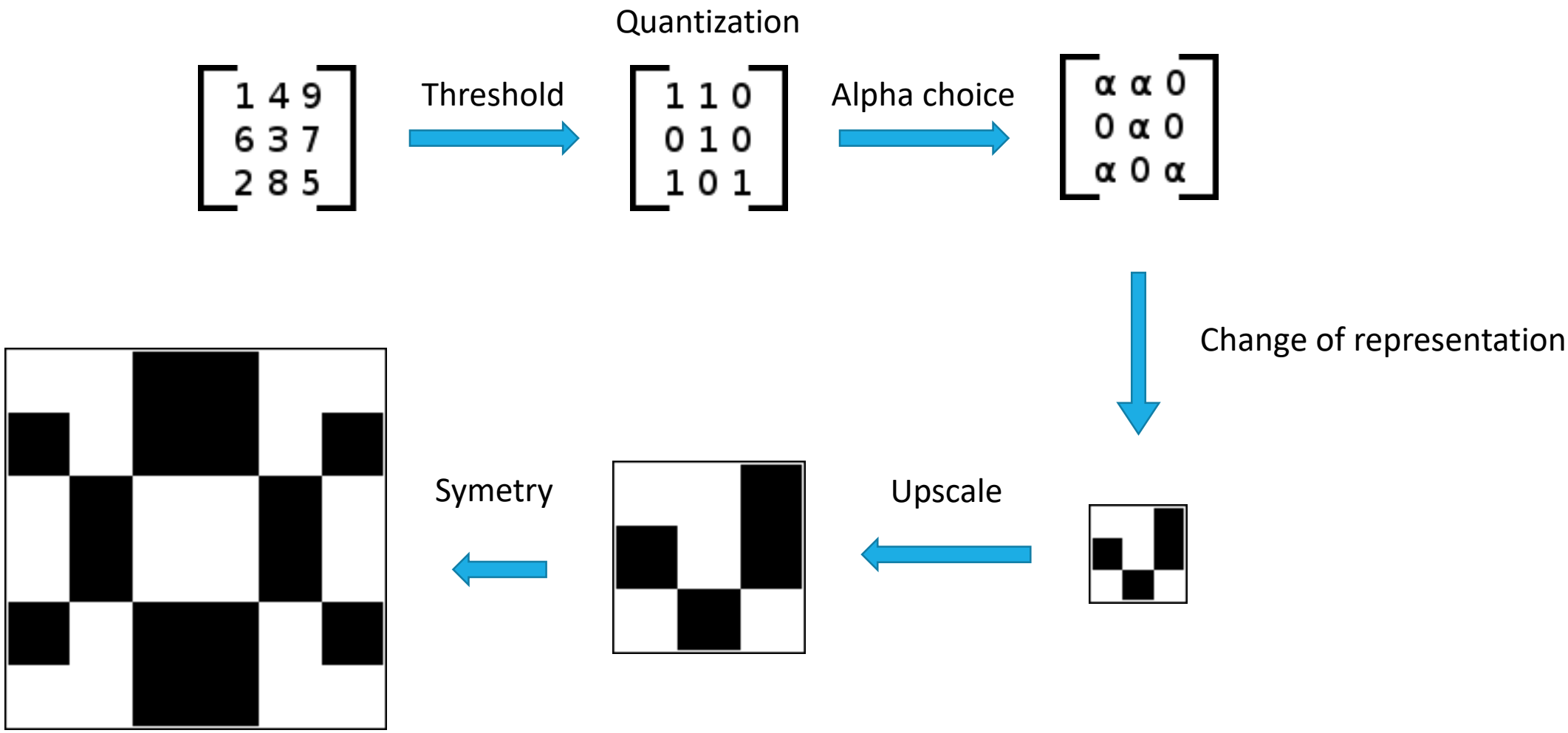
- Embed watermark in «textured» areas
- Create geometric fingerprint with specific characteristics
- Use the blue channel of the image to embed the watermark

Textured areas

Visible watermark

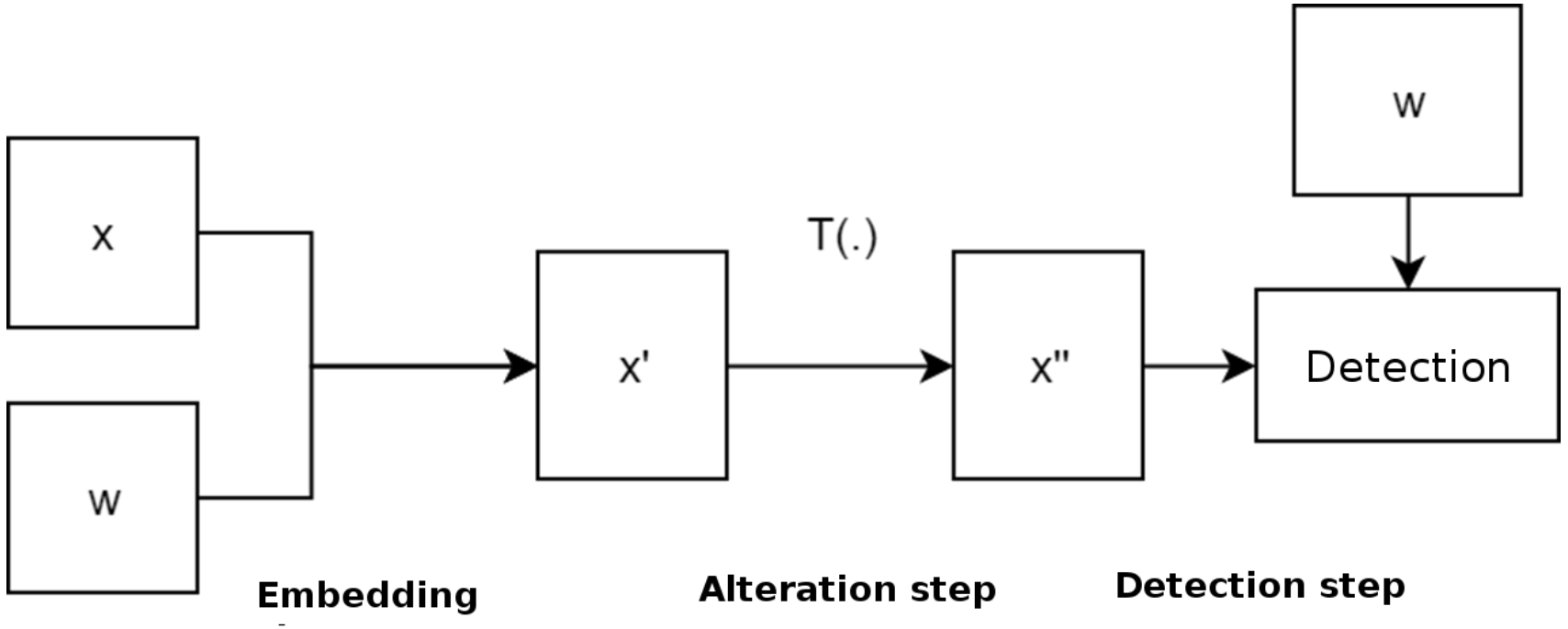
Invisible watermark

Watermark generation



Embeding

Watermark : $\boldsymbol{w} \in \mathbb{R}^{q \times q}$ Image : $\boldsymbol{x} \in \mathbb{R}^{m \times n}$



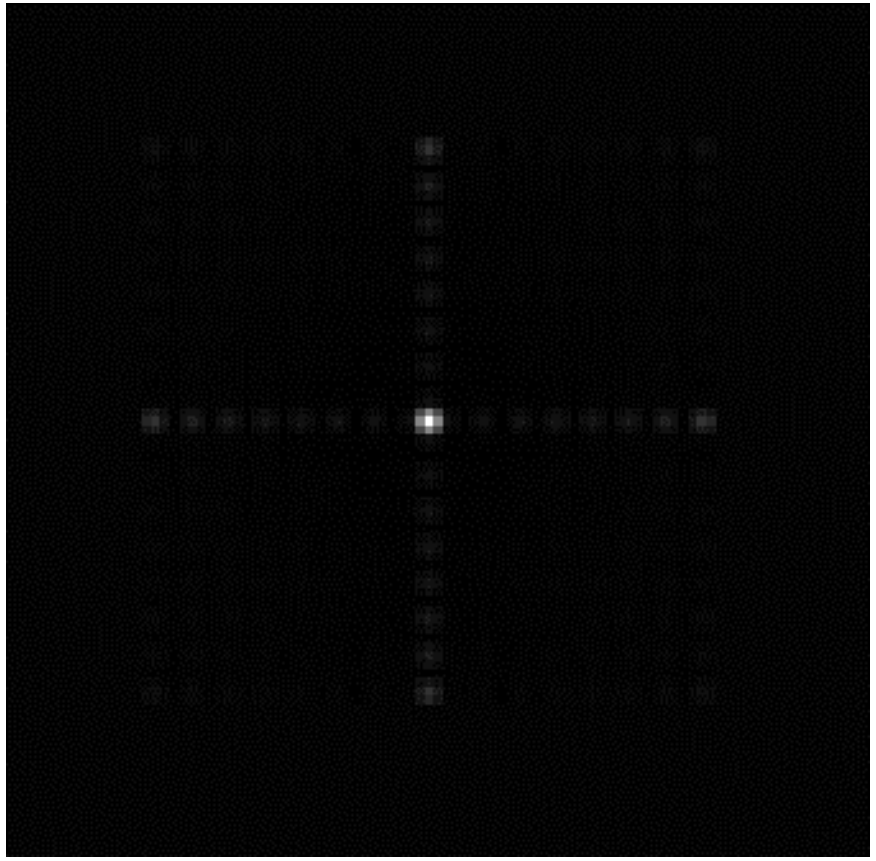
Watermark detection

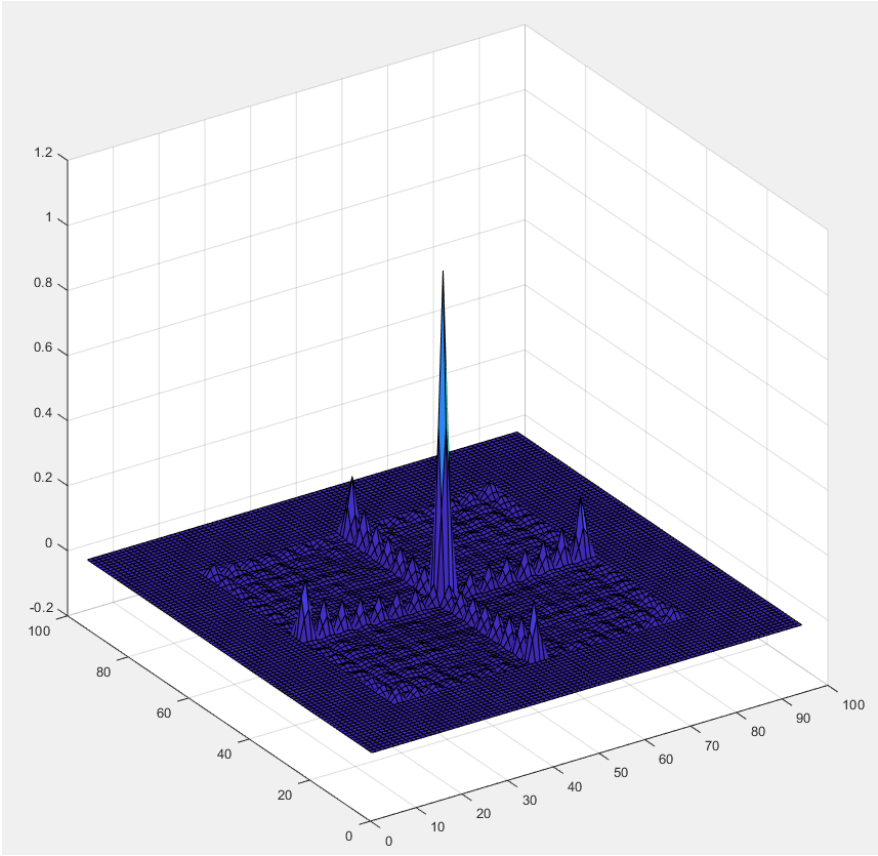
Residual image $\boldsymbol{I_r}$:

$$\mathrm{Blur}(\boldsymbol{x''}) = \boldsymbol{I}$$

$$\boldsymbol{I_r} = \boldsymbol{I} - \boldsymbol{x''}$$

Symmetry pattern





Efficiency

Transform domain methods

Robustness attack (unauthorized removal)

- Additive noise: This may happen (unintentionally) in certain applications such as printing and scanning converters or from transmission errors. It could happen intentionally by an attacker who is trying to destroy the watermark (or make it undetectable) by adding noise to the watermarked cover
- Filtering: Linear filtering such as low-pass filtering or non-linear filtering such as median filtering.
- Lossy compression: This is generally an unintentional attack, which appears very often in multimedia applications. The compression algorithm might unadvertently remove the watermark
- Collusion attack: In some watermarking schemes, if an image has been watermarked many times under different secret keys, it is possible to collect many such copies and “average” them into a composite image that closely resembles the original image and does not contain any useful watermarking data.

Presentation attack (masking attack)

- The goal is not to remove the watermark completely but to break the detection algorithm
- Cropping: This is a very common attack since in many cases the attacker is interested in a small portion of the watermarked object, such as parts of a certain picture or frames of a video sequence. With this in mind, in order to survive this kind of attack, the watermark needs to be spread over the dimensions where this attack takes place.
- Rotation and spatial scaling: Detection and extraction might fail when rotation or scaling is performed on the watermarked image

An Interpretation Attack

- The attacker is trying to create a situation that prevents assertion of ownership
- Multiple watermarking: An attacker may watermark an already watermarked object (creating uncertainty about which watermark was inserted first) and later make claims of ownership. The easiest solution is to timestamp the hidden information by a certification authority.
- Unauthorized embedding (forgery): Embed illegitimate watermark into works that should not contain them.

Advantages of the transform domain

- By transforming spatial data into another domain, statistical independence between pixels can be obtained
- The watermark is irregularly distributed over the entire spatial image upon inverse transform, which makes it more difficult for enemies to decode and read the mark.
- Cropping may be a serious threat to any spatially based watermark but is less likely to affect a frequency-based scheme. Since watermarks applied to the frequency domain will be dispersed over the entirely of the spatial image upon inverse transformation so we can retrieve part of the watermark.

HVS and frequency domain

- It is usually true that human eyes are not sensitive to small changes in edges and texture but they are very sensitive to small changes in the smooth areas of an image
- In other words, the human eyes are more sensitive to lower frequency noise, rather than high frequency noise.
- The watermark should be embedded into higher frequency components to achieve better perceptual invisibility, however, high frequencies might be discarded after most attacks such as lossy compression or shrinking
- In order to prevent the watermark from being easily attacked, it is often necessary to embed the watermark in the lower frequency coefficients. The attacker trying to change these coefficients is likely to significantly affect image quality. However, human eyes are more sensitive to lower frequency noise.

Thank you for your attention