DWT and QR-code based watermarking for document DRM

Nicolò Cardamone and Fabrizio d'Amore

Sapienza University of Rome, Italy
cardamone.1613937@studenti.uniroma1.it
damore@diag.uniroma1.it

Abstract. This paper presents a digital rights protection scheme for every type of document containing images or text using a number of steps that uses cryptography and watermarking. The entities involved in this process are two: the owner of the document that has digital rights on it and a generic user who can download or view the watermarked version of the original document. The watermarked document contains a QR-code that is repeatedly inserted, and scrambled, by the document right's owner, into the frequency components of the image, thus producing the watermarked image. The signed ID uniquely identifies every users using the system. The schema, a non-blind type, achieves good perceptive quality and fair robustness using the 3rd level of the Discrete Wavelet Transform. The experimental results show that, inserting more occurrences of a scrambled QR-code, the proposed algorithm is quite resistant to JPEG compression, rotation, cropping and salt and peeper noise.

Keywords: Watermarking · Steganography · DWT

1 Introduction

In a big company data protection is one of the most important thing due to the fact that data leaks about company strategies, assets or every other types of confidential documents could destroy the business of the entire company and its economic value and subsequently it can lead to bankruptcy. For this reason is useful to mark document that are given to employees or third party in order to know who is the responsible of the eventual leak. There exist an approach [2] that relies on DWT and QR-code based watermarking applied to coloured or black and white images. In this paper this approach will be adapted to document and we will show that the use of DWT, QR-code and cryptography assures good imperceptibility, watermark extraction performance and robustness against most common image manipulations like JPEG compression, rotation, cropping and additive noise. Every document can be seen as an image for example a pdf file of X pages can be converted in X images, some changes are needed because a document is made up mainly of black and white without shades of gray if it has no pictures inside but in some cases for example course slides with background are very similar to a classic picture and the approach [2] works with no problem.

The main idea is to embed data into the image's lower frequency of the 3rd level Discrete Wavelet components, this data is encrypted and uniquely identifies an person, if the document will be leaked we're able to extract this data, even if some modifications are made on the document, and understand who is the responsible of the leak. QR-code will be inserted as watermark because of its error correction capability, in order to improve imperceptibility and extraction performances, we insert it into the host image more times, in a key-scrambled version. This is a non-blind schema so to extract the watermark is necessary to provide the original image.

2 Preliminaries

Watermarking techniques can be classified into different categories according to the type of domain in which data embedding takes place and the type of information is needed to extract the watermark. There are mainly two domain types that are spatial and frequency. When we want an invisible watermark we usually use steganography techniques because we want to hide payload in a document in this way an user without a thorough analysis can't distinguish between a watermarked document and a non watermarked one, otherwise the watermark can be visible but in this case can be removed. Regarding what is needed to extract the watermark from the image we can divide the cases into blind, semi-blind and non-blind systems. A blind watermark, or public watermarking algorithm, requires neither the original image nor the embedded watermark to extract it from the watermarked image; a semi-blind, or semi-private scheme, requires only the watermark and finally a non-blind scheme requires at least the cover image. In this paper we will use a non-blind schema and only the cover image is needed to recover the original QR-code inserted. Spatial image watermarking techniques are commonly used in steganographic context because, hiding data into the least significant bits of an image, achieves to embed large quantity of data but the watermark is not robust to common manipulations like JPEG compression. In frequency domain we have two main techniques Discrete Cosine Transform (DCT) or Discrete Wavelet Transform (DWT) the first one is not resistant to rotation, translation and image cropping due to the block divide algorithm while the second one assures good robustness against the most popular image manipulations. Digital images properties can be better expressed through a wavelet transform since the frequency components are quickly varying around the image area. Through the wavelet decomposition the original signal can be represented by its coefficients which contains the spatial information. Each level of a DWT produces four types of coefficients: LL, or approximation coefficients, that represent the low frequency part of the image (most of information) and the details coefficients LH, HL and HH (vertical, horizontal and diagonal). In every level the decomposition is obtained on the LL component of the previous level. The original signal can be completely reconstructed performing the Inverse Wavelet Transformation on these coefficients.

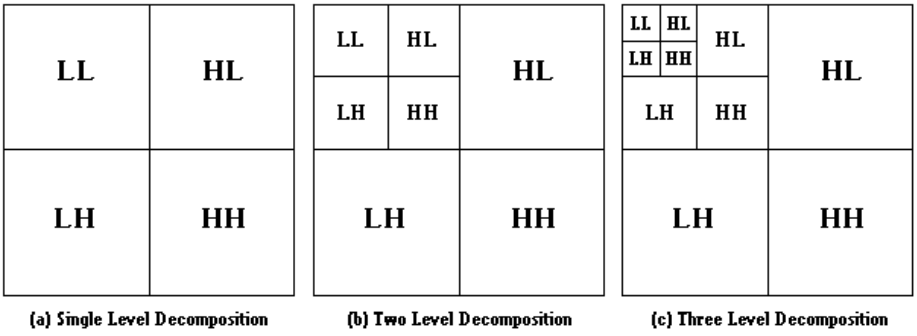


Fig. 1. DWT levels 1, 2 and 3.

In order to achieve a good visual imperceptibility, according to the spectral sensitivity of human eye, the blue component of a colour image is most suitable for hiding data and this component will be used in our approach. Data hiding system performances are described in terms of imperceptibility, embedding capacity and robustness. For digital watermarking the most important are imperceptibility and robustness. Some measures will be made on the final image to measure visual imperceptibility between the original image and the watermarked one that are Mean Squared Error (MSE), Peak Signal to Noise Ratio (PSNR) and Structural Similarity:

$$\text{MSE} = \frac{1}{m \times n} \sum_{i=1}^{m} \sum_{j=1}^{n} (X(i,j) - X'(i,j))^2$$

$$\text{PSNR} = 10 \log_{10} \frac{\text{MAX}_i^2}{\text{MSE}}$$

$$\text{SSIM} = \frac{(2\mu_i\mu_j + C_1)(2\sigma_{ij} + C_2)}{(\mu_i^2 + \mu_j^2 + C_1)(\sigma_i^2 + \sigma_j^2 + C_2)}$$

Where m and n are the number of rows and columns of the image expressed in pixel, $X(i,j)$ is the value of the pixel at row i and column j of the original image, $X'(i,j)$ is the value of the pixel at row i and column j of the watermarked image, MAX_i is the biggest value of a pixel, μ, σ, σ_{ij} are, respectively, mean, standard deviation and correlation, and C_1, C_2 are constants.

The payload is contained in QR code (Quick Response Code) that is the trademark for a type of matrix barcode first designed in 1994 for the automotive industry in Japan. A QR code consists of black squares arranged in a square grid on a white background, which can be read by an imaging device such as a camera, and processed using Reed–Solomon error correction until the image can be appropriately interpreted. The required data is then extracted from patterns that are present in both horizontal and vertical components of the image.

3 Approach

In this section are described all the steps of the process of the watermarking and the extraction of QR code from the watermarked image. Suppose all the documents that will be watermaked come from a single entity, a pair of keys public and private is created, and a message containing the ID of the user that will get the confidential document is signed by the entity using the private key. Starting from the original image and the signed message, the entity produces the watermarked image for the user by the following steps:

1. Convert each page of the document in an image.
2. Computes the approximation coefficients of level 3 ($\mathrm{AC_{LL3}}$) by performing a third-level decomposition of the image using a wavelet (blue component in case of colour image).
3. Generate a QR code encoding the hmac-sha256 of the signed message using his private key.
4. Derives a scrambling key from a hmac-sha256 of a password and use it to scramble the QR code repetitions necessary to fit the size of the $\mathrm{AC_{LL3}}$ (N $\times$ M) of the image obtaining WIM like in Fig. 2.
5. Insert the watermark into the approximation coefficients of level 3 of the watermarked image $\mathrm{WAC_{LL3}}(i,j) = \mathrm{AC_{LL3}}(i,j) + k \times \mathrm{WIM}(i,j)$, $i = 1, 2, \ldots, N$ and $j = 1, 2, \ldots, M$, with $k = 20000$ for colour images and $k = 15000$ for black and white ones.
6. Obtain the watermarked image by performing the inverse discrete wavelet transform.
7. Reconstruct the document starting from the watermarked images.

To extract the QR code starting from the original image (Fig. 1) and the watermarked image (Fig. 5 A) it is necessary to:

1. Convert each page of the document in an image.
2. Compute the approximation coefficients of level 3 by performing a third-level decomposition of the image using a wavelet, blue component in case of colour image, for both original ($\mathrm{AC_{LL3}}$) and watermarked image ($\mathrm{WAC_{LL3}}$).
3. Reconstruct WIM:

$$f(x) = \begin{cases} 1 & \text{if } \mathrm{WAC_{LL3}}(i,j) - \mathrm{AC_{LL3}}(i,j) \geq t \\ 0 & \text{otherwise} \end{cases}$$

   for $i = 1, 2, \ldots, N$ and $j = 1, 2, \ldots, M$, with $t = 40$ for colour images and $t = 22.5$ for black and white ones, obtaining a scrambled watermarked image.
4. Descramble by using the key derived from the hmac-sha256 of the previously generated password in the third step of watermark insertion.
5. Compute hmac-sha256 of the message using entity private key.
6. Recover the QR code from the single QR code repetitions occurring in the descrambled image (Fig. 5 B) and verify if the decoded value is equal to the hmac-sha256 of the message, for payload extraction is used either each single extracted repetition of a QR code either a reconstructed QR code based on majority pixel value matching, upon 1 to the maximum value of them.

The watermark extraction procedure can be done only by the entity because we need his private key and the unscrambling password, if an attacker finds the password he can't generate another document watermarked with another ID because he didn't know the private key of the entity, if he is able to obtain public and private keys of the entity he can not once again generate another document because he can generate an unscrambled watermarked version of the document like Fig. 5 B but he can't scramble QR code like Fig. 5 (a).

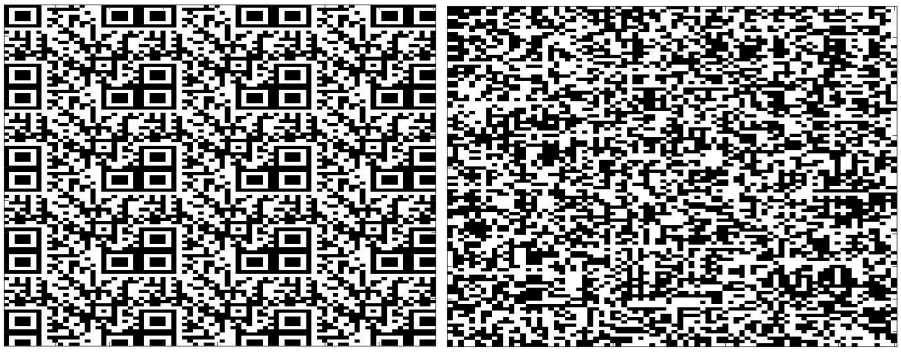


Fig. 2. Unscrambled and Scrambled QR code.

4 Applications

We implemented this technique in MATLAB using as original images the ones represented in Fig. 3, 4, 6, and 8. We used the open-source library "libqrencode" [3] for QR code generation and the "quirc" library [1] for QR code decoding. All images were tested with the first 20 wavelets of the Daubechies family. In this section we're going to apply the procedure to three class of documents:

- Black and White text document
- Text document with modified background
- Slides with background

4.1 Black and White text document

Black and white document needed a little preprocessing because some information of QR code were lost during image reconstruction from wavelet transform, the background is slightly darker as we can see in Fig. 3 and 4 in this way no informations were lost.

In the Fig. 5 we can see the final result of the watermark insertion algorithm that produces an image with visible artefacts in the background, this is an example of visible watermark and an user that receive a document in this state can read it but understand that maybe some extra information has been added.

4.2 Text document with modified background

In the Fig. 6 we can see the final result of the watermark insertion algorithm and the original image side by side, in this case the watermark is invisible because it is visually imperceptible and is very difficult to notice differences between the two images.

4.3 Slides with background

In the Fig. 8 we can see the final result of the watermark insertion algorithm and the original image one above the other, in this case the watermark is slightly visible because the colours used in the slide are lighter than the example before and we can observe some background noise in the image.

4.4 Results

Results of the three test cases above are very good the QR code is extracted without error as we can see from Fig. 9. In table 1 we can find MSE, PSNR, SSIM results for the three types of document.

Table 1. MSE PSNR SSIM results.

Document type	MSE	PSNR	SSIM
Black and White text document	22.7515	79.5790	0.9215
Text document with modified background	47.8273	72.1493	0.8208
Slides with background	47.7314	72.1694	0.9392

4.5 Possible Attacks

There are some image manipulation attacks that can be performed on the watermarked image depending on the different types of document. In black and white document case an attacker can simply analyse pixel by pixel the document, if a pixel is black or white no modification is necessary, otherwise, as we can see in Fig. 5, convert the pixels in white one obtaining a result like in Fig. 3. In the slides case an attacker can do a similar thing like converting all the shade of a particular colour, in the example mainly red and blue, in a unique colour and this will result in failure in watermark extraction.

5 Final Remarks and Future Works

Through these steps we can embed scrambled individual references within an document with almost no effect on its quality. Experimental results show that such a schema provides quite good quality and robustness and the results show that the algorithm achieves fairly good results in terms of imperceptibility. Further analysis could be done by searching for new possibilities to insert information in a document for examples using fonts modification or introducing errors following a certain pattern.

References

1. Beer, D.: Quirc. `https://github.com/dlbeer/quirc`
2. Chiavarelli, S., d'Amore, F.: A novel approach to image DRM relying on DWT and QR-code based watermarking. Submitted (2018)
3. Fukuchi, K.: libqrencode. `https://fukuchi.org/works/qrencode`

DWT and QR-code based watermarking used for Document DRM

Nicolò Cardamone[1][0000−0002−1908−8247] and Fabrizio d'Amore[1]

University of Rome La Sapienza, Rome, Italy
cardamone.1613937@studenti.uniroma1.it
damore@dis.uniroma1.it

Abstract. This paper presents a digital rights protection scheme for every type of document containing images or text using a number of steps that uses cryptography and watermarking. The entities involved in this process are two: the owner of the document that has digital rights on it and a generic user who can download or view the watermarked version of the original document. The watermarked document contains a QR-code that is repeatedly inserted, and scrambled, by the document right's owner, into the frequency components of the image, thus producing the watermarked image. The signed ID uniquely identifies every users using the system. The schema, a non-blind type, achieves good perceptive quality and fair robustness using the 3rd level of the Discrete Wavelet Transform. The experimental results show that, inserting more occurrences of a scrambled QR-code, the proposed algorithm is quite resistant to JPEG compression, rotation, cropping and salt and peeper noise.

Keywords: Watermarking · Steganography · DWT

1 Introduction

In a big company data protection is one of the most important thing due to the fact that data leaks about company strategies, assets or every other types of confidential documents could destroy the business of the entire company and its economic value and subsequently it can lead to bankruptcy. For this reason is useful to mark document that are given to employees or third party in order to know who is the responsible of the leak. There exist an approach [1] that relies on DWT and QR-code based watermarking applied to coloured or black and white images. In this paper this approach will be adapted to document and we will show that the use of DWT, QR-code and cryptography assures good imperceptibility, watermark extraction performance and robustness against most common image manipulations like JPEG compression, rotation, cropping and additive noise. Every document can be seen as an image for example a pdf file of X pages can be converted in X images, some changes are needed because a document is made up mainly of black and white without shades of gray if it has no pictures inside but in some cases for example course slides with background are very similar to a classic B/W picture and the approach [1] works with no problem. The main

Fig. 3. W/B preprocessed text example.

DWT and QR-code based watermarking used for Document DRM

Nicolò Cardamone[1][0000−0002−1908−8247] and Fabrizio d'Amore[1]

University of Rome La Sapienza, Rome, Italy
cardamone.1613937@studenti.uniroma1.it
damore@dis.uniroma1.it

Abstract. This paper presents a digital rights protection scheme for every type of document containing images or text using a number of steps that uses cryptography and watermarking. The entities involved in this process are two: the owner of the document that has digital rights on it and a generic user who can download or view the watermarked version of the original document. The watermarked document contains a QR-code that is repeatedly inserted, and scrambled, by the document right's owner, into the frequency components of the image, thus producing the watermarked image. The signed ID uniquely identifies every users using the system. The schema, a non-blind type, achieves good perceptive quality and fair robustness using the 3rd level of the Discrete Wavelet Transform. The experimental results show that, inserting more occurrences of a scrambled QR-code, the proposed algorithm is quite resistant to JPEG compression, rotation, cropping and salt and peeper noise.

Keywords: Watermarking · Steganography · DWT

1 Introduction

In a big company data protection is one of the most important thing due to the fact that data leaks about company strategies, assets or every other types of confidential documents could destroy the business of the entire company and its economic value and subsequently it can lead to bankruptcy. For this reason is useful to mark document that are given to employees or third party in order to know who is the responsible of the eventual leak. There exist an approach [3] that relies on DWT and QR-code based watermarking applied to coloured or black and white images. In this paper this approach will be adapted to document and we will show that the use of DWT, QR-code and cryptography assures good imperceptibility, watermark extraction performance and robustness against most common image manipulations like JPEG compression, rotation, cropping and additive noise. Every document can be seen as an image for example a pdf file of X pages can be converted in X images, some changes are needed because a document is made up mainly of black and white without shades of gray if it has no pictures inside but in some cases for example course slides with background are very similar to a classic picture and the approach [3] works with no problem.

Fig. 4. W/B postprocessed text example.

Fig. 5. Scrambled and Unscrambled Watermarked image.

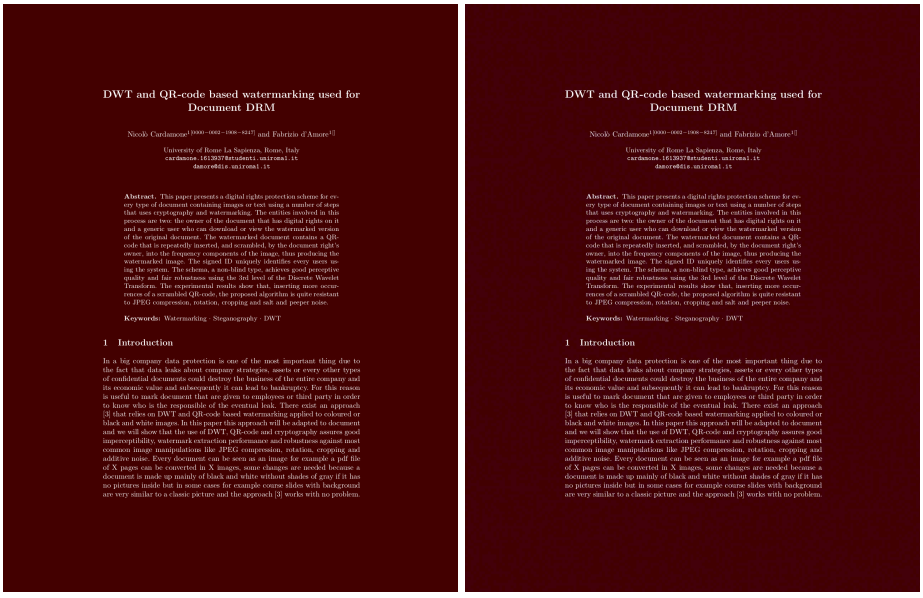


Fig. 6. Side by side comparison between original image and watermarked one.

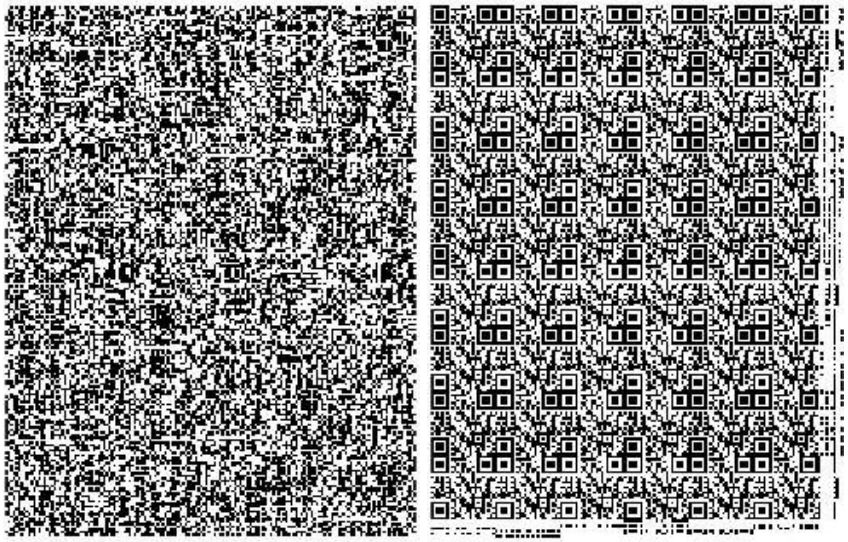
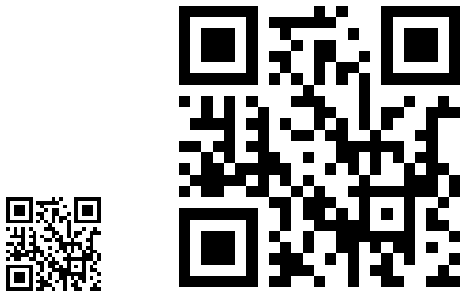


Fig. 7. Scrambled and Unscrambled QR code extracted.

(a)

(b)

Fig. 8. (a) The original slide; (b) the watermarked slide.



(a) (b)

Fig. 9. QR-code: (a) inserted into the image; (b) extracted side by side.