Exploitation of redundancy for pattern estimation of copy-sensitive two level QR code

Iuliia Tkachenko, Christophe Destruel

Exploitation of redundancy for pattern estimation of copy-sensitive two level QR code

Iuliia Tkachenko
Laboratory Hubert Curien, UMR CNRS 5516
University Lyon, UJM-Saint-Etienne
18 Rue Professeur Benoît Lauras
42000 Saint-Etienne, France
iuliia.tkachenko@univ-st-etienne.fr

Christophe Destruel
IGO
78, rue John Mac Adam
30900 Nimes, France
christophe.destuel@igo.fr

Abstract

The falsification of hardcopy documents is a common problem these days. Numerous industrial and scientific solutions have been proposed to prevent these falsifications. In this paper, we want to study the security of the two level QR code which is constructed using specific textured patterns that are sensitive to print-and-scan impact. Such code is a good candidate because it generalizes several concepts from several codes. We take a falsifier point of view that aims to reconstruct the two level QR code and to fool the authentication system detector. As the two level QR code contains sets of the same textured patterns, the opponent has access to different printed-and-scanned versions of these textured patterns. These sets of patterns can be used for structure estimation. Several local strategies for pattern estimation are suggest in this paper. The experimental results show that the increasing number of printed-and-scanned patterns cannot improve the estimation results.

1. Introduction

The number of hardcopy documents that are processed by administration and authority centers increases every day and these documents cannot always be replaced by their numeric version (for example, the validity of an ID document is also linked to the physical object). In the same time the number of administrative document falsification increases with accessibility of qualitative reproduction devices (such as high resolution printers and scanners). Therefore, the protection of administrative documents is a hot topic. Numerous solutions have been proposed: watermarking, text hash calculation [1] and copy sensitive codes [2].

Copy Sensitive Graphical Codes (CSGC) are a promising solution for document support authentication. These codes have been constructed to be sensitive to distortions added by Print-and-Scan (P&S) process during each document reproduction. The P&S process is a stochastic process which introduces changes that cannot be reproduced [3]. One of the first commercialized CSGC was the Copy Detectable Pattern (CDP) [2]. This is a black-and-white maximum entropy image, generated using a secret key, that takes full advantage of information loss principle during P&S process. The authentication test performs the comparison of the pixel values in the digital and in the scanned CDPs. This solution resists to some attacks (duplication, estimation using inverse P&S model [4]). However, it was shown that an attacker can produce a fake that fools the detector with reasonable number of genuine goods [5].

Another CSGC is the Two Level QR (2LQR) code introduced in [6]. This barcode stores the information in two levels. The first level is readable by any standard barcode reader, the second level can be accessible only for authorized users who possess the specific reading application. Additionally, the second level is sensitive to P&S process. The second level, as well as the authentication capacity, is offered by specific texture patterns, that are distinguished one from another after P&S process and can be sensitive to coping process. The authentication test is based on comparison of mean 2LQR code correlation value with predetermined threshold Th. The document is said to be authentic if this mean value is bigger than Th. It was shown that the 2LQR code can successfully resist to duplication attack [7]. Additionally, the 2LQR code can resist several naive attacks based on the use of some image processing tools (histogram equalization, global image binarization and sharpen enhancement) [8].

The 2LQR code second level encodes the information using a q-ary ($q \geq 2$) alphabet, which is composed of q different textured patterns. The structure and the number q of these textured patterns are kept as a secret for further authentication test but the limited number of letters in an alphabet can

be seen as a weakness. Due to this specific construction, we can imagine a pattern estimation attack based on a big number of samples: an opponent will use a big amount of the same pattern (that can be extracted from a single code) to estimate the original texture and then to reproduce the document with this original-like graphical code. In this paper, we want to discuss and suggest several basic opponent strategies to forge the 2LQR code.

The rest of the paper is organized as follows. We introduce the CSGC construction and the authentication process challenge in Section 2. The suggested pattern estimation methodology is discussed in Section 3. The binarization strategies for textured pattern estimation are listed in Section 3.2. We show and analyze the experimental results in Section 4. Finally, we conclude in Section 5.

2. Challenge statement

In this section, we want to explain the specificity of the 2LQR code construction and introduce the authentication and estimation processes from theoretical point of view.

2.1. Graphical code construction

As it was mentioned before the supplementary information of 2LQR code is encoded using q different textured patterns that are chosen according to some criteria introduced in [6]. A specificity of the 2LQR code is to propose a new data storage capacity linked to its sensitivity to P&S process. These functionalities are based on the use of a finite q-ary alphabet of textures: by definition the letters of the alphabet are repeated and this repetition offers a new attack way to opponents. During authentication test, we compare each printed-and-scanned textured pattern (i.e. modules of 2LQR code) with original textured patterns and we test the response of each of them, but an opponent can try to take advantage of this pattern repetition to estimate globally the textured patterns used in order to create an unauthorized copy of a given 2LQR code.

The difference between CDP and textured patterns of 2LQR code is that the CDP is composed of black-and-white pixel squares that can be considered one by one (to store binary information) or globally (to ensure support authenticity). This one-by-one method to store information impacts the precision of the devices used and or the reliability of the process. On the other hand the elementary structure of a 2LQR code is a textured module in which several pixels are either black or white. This textured module is used for both functions at the same time: data storage and document authentication.

Let U be a square image of size $u \times u$ pixels that is the elementary piece of CDP image. Let P be a textured pattern of size $p \times p$ pixels that represents the elementary piece of 2LQR code. We illustrate the comparison of elementary pieces of CDP and 2LQR code in Fig. 1. The smallest element of CDP image is an image U that is either black or white, so all u^2 pixels are equal to 0 or 255. In the same time, the smallest element of 2LQR code is a textured pattern that contains a frequent change of black (0) and white (255) pixels. Thus the textured patterns have a more detailed structure than the CDP elements.

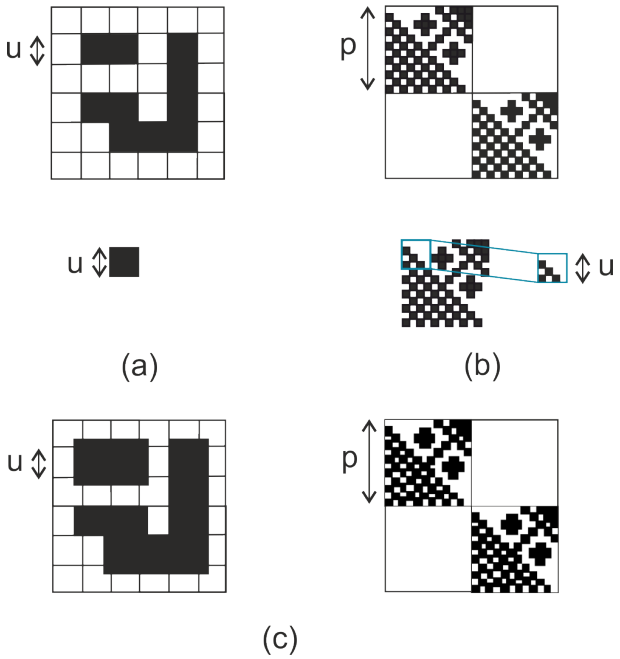
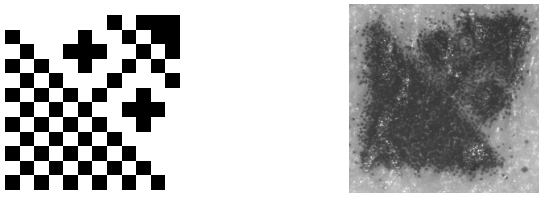


Figure 1. The comparison of smallest elements of a) CDP and b) 2LQR code, and c) the impact of P&S process to these CSGC.

As shown in Fig. 1.c, due to P&S impact each element U of CDP is spread over $\Lambda \times \Lambda$ area, where $\Lambda = \lambda \times u$. In the same time, each pixel of textured pattern P is influenced by its neighbors so that a printed and scanned textured pattern is spread over $s \times s$ area, where $s = \omega \times p$. However, in practice, the value s is often equal to p or bigger just up to one pixel. In this paper, we consider that $s = p$ or $s = 2 \times p$ when the scanner resolution is twice bigger than the printer resolution.



(a) (b)

Figure 2. Example of textured pattern: a) original version and b) its printed version captured using ZEISS microscope with $5\times$ magnification.

In order to illustrate the crucial impact of P&S process on textured patterns used in 2LQR code, we print one textured pattern using common office printer with 600 dpi resolution

and magnified the printed pattern using ZEISS microscope ($5\times$ magnification). Fig. 2.b illustrates the obtained image that shows a noise added to each pixel of textured pattern. We note that even with high resolution microscope we cannot correctly predict the original texture of pattern.

2.2. Authentication process

The authentication is based on the comparison of original CSGC image $\mathbf{I}$ with its degraded by P&S process CSGC version $\tilde{\mathbf{I}} = \mathbf{I} + \mathbf{N}_P + \mathbf{N}_S$, where $\mathbf{N}_P$ and $\mathbf{N}_S$ are noises added by authorized printer and scanner. The duplication attack consists on two consecutive P&S operations so that the duplicated CSGC image will be

$$\tilde{\tilde{\mathbf{I}}} = \mathbf{I} + \mathbf{N}_P + \mathbf{N}_S + \mathbf{N}_P + \mathbf{N}_S,$$

when we suppose that the opponent has the same printer and scanner as the authority center.
The authentication test measures the quantity of distortion added to printed CSGC $\tilde{\mathbf{I}}$ in order to take a decision about its authenticity. Therefore the authentication test can be formulated as hypothesis test:

$$H_0 : f(\mathbf{I}, \tilde{\mathbf{I}}) > Th,$$
$$H_1 : f(\mathbf{I}, \tilde{\mathbf{I}}) < Th,$$

where f is any comparison function as a distance between the images or a correlation value and Th is the distortion threshold that was calculated in advance by the authority center. The hypothesis H_0 is valid when the printed CSGC $\tilde{\mathbf{I}}$ comes from normal mode and thus is authentic. The hypothesis H_1 is valid when the original CSGC image was duplicated ($\tilde{\mathbf{I}} = \tilde{\tilde{\mathbf{I}}}$) or faked.
In the case of estimation attack, an opponent scans an original printed CSGC $\tilde{\mathbf{I}}$ and then tries to estimate the CSGC image $\hat{\mathbf{I}}$ that will be than equivalent to the original $\mathbf{I}$. In this case, the fake CSGC can probably pass the authentication test as:

$$\tilde{\tilde{\mathbf{I}}} = \hat{\mathbf{I}} + \mathbf{N}_P + \mathbf{N}_S.$$

It was shown in [9] that an opponent can estimate the structure of CDP and produce a fake with reasonable number of genuine CDP samples that will be considered as an original by the authenticator. In this paper, we want to test the close local strategies to estimate the structure of each textured pattern used for 2LQR code construction.

3. Estimation of copy-sensitive 2LQR code

The textured pattern redundancy in 2LQR code can make easier its estimation. We suppose that an opponent can easily predict the dimension q used in 2LQR code, and then s/he can use a classification test for textured pattern classification. These classes can be used then for pattern estimation. In this section we discuss several possible local strategies for pattern estimation.

3.1. Collection of samples from 2LQR code

We suppose that the opponent has access to one or several P&S samples of a 2LQR code and does not have the original numeric patterns used for its generation. In this case, the opponent needs to reconstruct each textured pattern. For this s/he needs to extract the P&S samples of textured patterns that constitute the given P&S 2LQR code. After, based on this dataset s/he can try to estimate the structure of these textured patterns using several binarization techniques. Therefore, this falsification strategy consists of textured pattern estimation using a big number of samples. The flowchart of proposed estimation process is presented in Fig. 3
Remark: we focus on pattern binarization because this process is a key step in the printing process and can not be avoid. Toner of common printers only deals with black and white information: greyscale data is dithered by the printer itself.

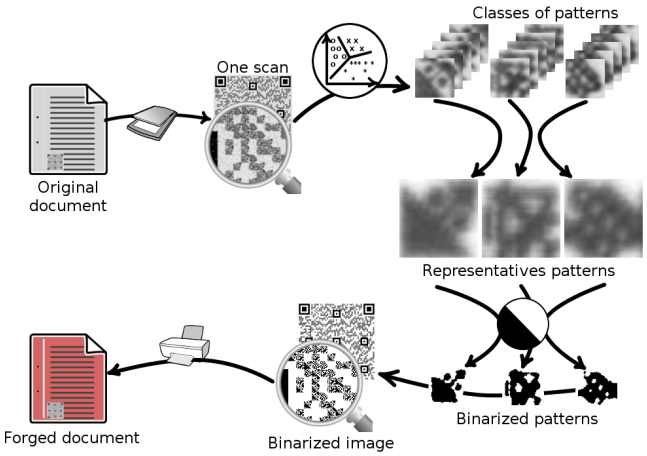


Figure 3. Considered estimation process. The patterns from P&S 2LQR code are classified in several classes. Then representative candidates (for each class) are created by estimation and binarization methods. Finally, a fake 2LQR code is constructed and printed by an opponent.

We suggest to use a k-mean clustering for pattern classification to q classes, where q is the dimension of the alphabet used for private level encoding in the given 2LQR code. *Remark: we use the k-means clustering as it is an effective classifier for our experiments and produces good classifications. Nevertheless, it can be replaced by any other classification/clustering method.* After separation of textured patterns into q classes, one of the binarization methods, discussed in following section, is applied for each class. Even if these binarization methods will not reconstruct the exact structure of the original textured patterns, the goal is to pass the authentication test currently based on Pearson correlation. So an opponent wants to respect the H_0 hypothesis:

$$cor(I, \tilde{\hat{I}}) > Th,$$

where $\hat{I}$ is estimated 2LQR code with q estimated patters $E_i, i = 1, \cdots, q$.

3.2. Estimation methods

In this section we overview several binarization methods to estimate the textured patterns after P&S process.

Let $S_i, i = 1, \cdots, q$ be a printed-and-scanned version of textured pattern P_i. Let $Class_i, i = 1, \cdots, q$ be the set of N_i textured patterns $S_i^j, j = 1, \cdots, N_i$ obtained after k-mean clustering, M_i be the mean image of size $s \times s$ pixels calculated using a set of N_i patterns S_i^j from $Class_i$ and E_i be an estimated binary pattern of size $p \times p$ pixels that might have the texture closed to P_i. In the rest of this section, we use these sets of patterns for pattern estimation.

In **the first type** of methods, we binarize the image sets $Class_i$ using one of the following thresholds:

- B1 method: $th = (min(Class_i) + max(Class_i))/2$;
- B2 method: $th = 127$;
- B3 method: $th = median(Class_i)$.

Then to construct E_i, the majority vote is applied for each pixel position in textured pattern in order to decide if this pixel is black or white.

In **the second type** of methods, we first calculate the mean image (M_i) using each image from the $Class_i$ set and then, each mean pattern M_i is binarized using one of the three following thresholds:

- B4 method: $th_i = mean(Class_i)$ - the mean value of each image set is used as a th;
- B5 method: $th_i = mean(M_i)$ - the mean value of calculated mean image is used as a th;
- B6 method: $th = 127$.

In the end we obtain the resulting binary pattern E_i.

However, the considered global thresholding methods cannot correctly estimate the structure of textured pattern used as the pixel values are influenced by its neighbor pixels (see Fig. 1.c and Fig. 2.b). Therefore, we suggest methods that take into account the frequency of pixel values and the relations between 4-connected pixels.

The third type (B7) of methods binarizes the image sets using median value. Then for all pixels of one type (white or black), we calculate the sum of normalized pixel values. More precisely, we normalize each element of the class $Class_i$ in the interval $[0, 1]$, and then we calculate separately the sum of normalized values of white (black) pixels $S_w(m, n)$ ($S_b(m, n)$) using the binarization results, where (m, n) is the pixel from textured pattern. Then if $S_w(m, n) > S_b(m, n)$, the pixel (m, n) of pattern E_i is considered as a white pixel, otherwise it is a black pixel.

Algorithm 1 B7 binarization

Require: $Class_i, i = 1, \ldots, q$
1: Calculate $th = median(Class_i)$
2: $BClass_i$ is a binarization of $Class_i$ using th
3: **for** each pixel $(m, n) \in E_i$ **do**
4: $p = norm_p(E_i(m, n))$
5: **if** $BClass_i^j(m, n)$ is white **then**
6: $S_w(m, n) = S_w(m, n) + p$
7: **else**
8: $S_b(m, n) = S_b(m, n) + (1 - p)$
9: **end if**
10: **end for**
11: **for** each pixel $(m, n) \in E_i$ **do**
12: **if** $S_w(m, n) > S_b(m, n)$ **then**
13: pixel (m, n) is white
14: **else**
15: pixel (m, n) is black
16: **end if**
17: **end for**

The algorithm of this estimation method is shown in Algorithm 1.

The fourth type (B8) of methods is a modified version of the previous one. Here, we consider not only the pixel but its 4-connected neighbors; the sum of pixel values is replaced by the sum of pixel values of its 4-connected pixels. Then these values are also normalized in the interval $[0, 1]$ and the sum of normalized values of white (black) pixels $S_w(m, n)$ ($S_b(m, n)$) using the binarization results is calculated. Finally, the pixel of E_i is white, if $S_w(m, n) > S_b(m, n)$, otherwise the pixel is black.

4. Experimental results

In this section, we describe the database of 2LQR code used, discuss the pattern estimation results and show the authentication test results. Additionally, we present results considering changes in P&S patterns used.

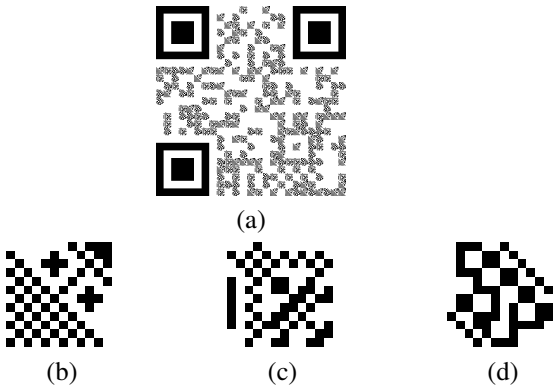


Figure 4. Examples of 2LQR code and textured patterns: a) 2LQR code, b) Pattern 1, c) Pattern 2, d) Pattern 3.

In our experiments, a version 2 of QR code (25×25 patterns) is used. The alphabet dimension of private level is $q = 3$. Therefore, we have more or less 85 textured patterns of each type in every 2LQR code. The database is composed of 30 P&S 2LQR codes. All 2LQR codes were printed with a Brother HL-4150CDN printer and scanned with a Canon LIDE210 scanner both with true 600 dpi black and white resolution. We set such resolution due to its common use for office printers. Fig. 4 illustrates a 2LQR code and textured patterns used.
As it was mentioned before the structure of textured pattern is crucially changed after P&S process. An example of each textured patterns after P&S is shown in Fig. 5. We can notice that the initial structure of textured patterns can hardly be recognized from these images.

(a) (b) (c)

Figure 5. Examples of P&S 2LQR code and textured patterns: a) P&S pattern 1, b) P&S pattern 2, c) P&S pattern 3.

We apply the suggested binarization methods for three classes of textured patterns. We notice that the frequent change of black and white pixels cannot be reconstructed by the proposed binarization methods after P&S process. The best reconstruction results can be obtained using method B8. We illustrate the comparison of numerical original pattern P_1, its magnified view after printing process using a microscope ZEISS, its printed-and-scanned version S_1 and its estimated version using B8 method in Fig. 6.

(a) (b) (c) (d)

Figure 6. Examples of : a) Pattern P_1, b) Pattern P_1 printed in 600dpi and captured using microscope, c) Printed-and-scanned pattern S_1, d) Estimated using B8 method pattern E_1.

4.1. Authentication results

As the final goal of the opponent is to fool the fake detection process, we use the authentication test as an indicator of reconstruction quality. The binarized patterns are used to regenerate the fake 2LQR codes, that are then printed using the same printer and scanner in 600 dpi resolution. These codes as well as the original 2LQR codes are tested to evaluate their authenticity. The correlation values for each 2LQR code (original or fake) are calculated. These correlation results for 30 samples are illustrated in Fig. 7. We note that the fake 2LQR codes have significantly smaller correlation values $(0.15 - 0.25)$ in comparison with originals that have in average correlation value equals to 0.35. With a classic authentication threshold set to $Th = 0.28-0.3$, no fake 2LQR code can pass the authentication test. All unauthorized duplication and falsifications of documents are detected (false positive rate is zero). Taking into account the fact that

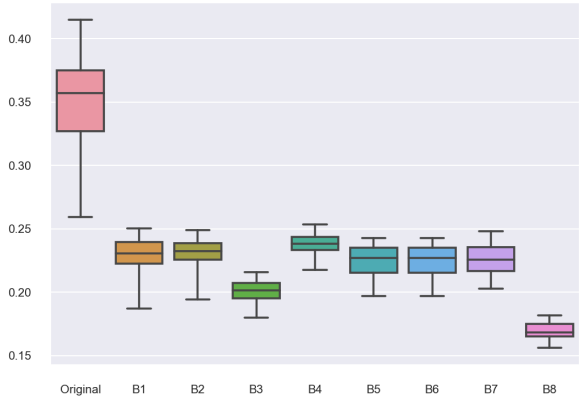


Figure 7. Comparison of correlation values for faked 2LQR code printed-and-scanned at 600 dpi resolution.

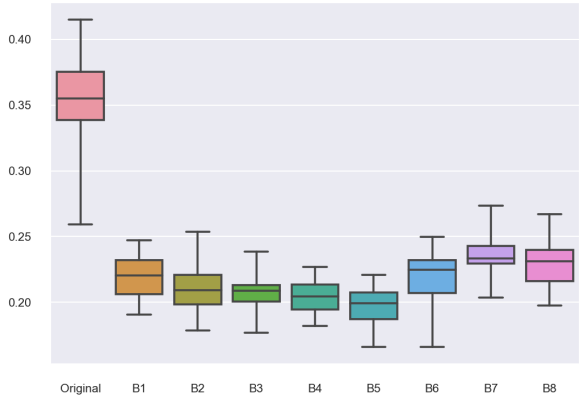


Figure 8. Comparison of correlation values for faked 2LQR code printed-and-scanned at 1200 dpi resolution.

the opponent can use higher resolution for fake code generation, we drive a test with 2LQR code printed-and-scanned with true one-channel 1200 dpi resolution. The results show that using specific attacks (B6-B8 for example) an opponent can reproduce a 2LQR code without destructing the stored information that remain readable. On the other hand, increasing the resolution is not enough to fake the anti-copy

test: the reproduced code is still detected as a copy (see Fig. 8).

4.2. Estimation results vs number of samples used

In this section we want to show some results of pattern estimation depending on code number used for this estimation. We construct several fake 2LQR codes in order to verify the authentication test robustness against these attacks. In the first experiment, three and ten samples of 2LQR codes were used for estimation of textured patterns using B8 method. These samples were printed and scanned using the same devices in 600 dpi resolution. The correlation values of these two estimations for 30 fake 2LQR codes are shown in Fig. 9, violet and yellow boxes. We note that the increased number of 2LQR code samples does not improve the estimation results. The correlation values are almost the same and thus the fake codes cannot pass the authentication test. Analogically, we change the number of 2LQR code samples from 3 to 30 for pattern estimation and use the B3 binarization method. The correlation values of these two sets of fake codes are illustrated in Fig. 9, green and red boxes. As in previous example, increasing the number of 2LQR codes cannot improve the textured pattern estimation results. The correlation values are also close to each other and cannot help to pass the authentication test with pre-defined threshold.

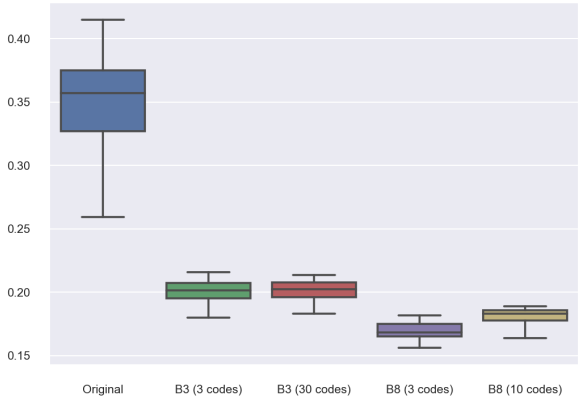


Figure 9. Comparison of correlation values, when changing the number of 2LQR codes.

5. Conclusions

In this paper, we suggest several opponent strategies to estimate the textured pattern structure and we evaluate the resistance of the 2LQR code. These methods are based on statistical values, majority vote and 4-connected pixels. The results estimated on many codes show that the initial structure of the textured patterns used cannot be reconstructed using such binarization methods. The obtained results show that the constructed fake 2LQR codes cannot pass the authentication test even after a partial reconstruction codes. Increase the number of samples does not seem to improve the quality of the reconstructed patterns: noise consideration may not be sufficient. In future work we want to find the breaking point in the authentication process of a 2LQR code: finding such a limit will allow us to ensure an authentication capacity considering initial conditions. With this same goal, we want to study another security aspects of 2LQR codes by increasing the image resolution with super-resolution methods and by using more precise estimations based on optimization methods.

References

[1] R. Villán, S. Voloshynovskiy, O. Koval, F. Deguillaume, and T. Pun, "Tamper-proofing of electronic and printed text documents via robust hashing and data-hiding," in *Electronic Imaging 2007*. International Society for Optics and Photonics, 2007, pp. 65 051T–65 051T.

[2] J. Picard, "Digital authentication with copy-detection patterns," in *Electronic Imaging 2004*. International Society for Optics and Photonics, 2004, pp. 176–183.

[3] A. T. P. Ho, B. A. M. Hoang, W. Sawaya, and P. Bas, "Document authentication using graphical codes: Reliable performance analysis and channel optimization," *EURASIP Journal on Information Security*, vol. 2014, no. 1, p. 9, 2014.

[4] A. E. Dirik and B. Haas, "Copy detection pattern-based document protection for variable media," *Image Processing, IET*, vol. 6, no. 8, pp. 1102–1113, 2012.

[5] C. Baras and F. Cayre, "2D bar-codes for authentication: A security approach," in *Signal Processing Conference (EUSIPCO), Proceedings of the 20th European*, 2012, pp. 1760–1766.

[6] I. Tkachenko, W. Puech, C. Destruel, O. Strauss, J.-M. Gaudin, and C. Guichard, "Two-level QR code for private message sharing and document authentication," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 3, pp. 571–583, 2016.

[7] I. Tkachenko, W. Puech, O. Strauss, C. Destruel, and J.-M. Gaudin, "Printed document authentication using two level QR code," in *Acoustics, Speech and Signal Processing (ICASSP), 2016 IEEE International Conference on*. IEEE, 2016, pp. 2149–2153.

[8] I. Tkachenko, C. Destruel, O. Strauss, and W. Puech, "Sensitivity of different correlation measures to print-and-scan process," *Electronic Imaging*, vol. 2017, 2017.

[9] C. Baras and F. Cayre, "Towards a realistic channel model for security analysis of authentication using graphical codes," in *Information Forensics and Security (WIFS), 2013 IEEE International Workshop on*, 2013, pp. 115–119.