Two-Level QR Code for Private Message Sharing and Document Authentication

Iuliia Tkachenko, William Puech, *Senior Member, IEEE*, Christophe Destruel, Olivier Strauss, Jean-Marc Gaudin, and Christian Guichard

Abstract—The quick response (QR) code was designed for storage information and high-speed reading applications. In this paper, we present a new rich QR code that has two storage levels and can be used for document authentication. This new rich QR code, named two-level QR code, has public and private storage levels. The public level is the same as the standard QR code storage level; therefore, it is readable by any classical QR code application. The private level is constructed by replacing the black modules by specific textured patterns. It consists of information encoded using q-ary code with an error correction capacity. This allows us not only to increase the storage capacity of the QR code, but also to distinguish the original document from a copy. This authentication is due to the sensitivity of the used patterns to the print-and-scan (P&S) process. The pattern recognition method that we use to read the second-level information can be used both in a private message sharing and in an authentication scenario. It is based on maximizing the correlation values between P&S degraded patterns and reference patterns. The storage capacity can be significantly improved by increasing the code alphabet q or by increasing the textured pattern size. The experimental results show a perfect restoration of private information. It also highlights the possibility of using this new rich QR code for document authentication.

Index Terms—QR code, two storage levels, private message, document authentication, pattern recognition, print-and-scan process.

I. INTRODUCTION

TODAY, graphical codes, such as EAN-13 barcode [1], Quick Response (QR) code [3], DataMatrix [2], PDF417, are frequently used in our daily lives. These codes have a huge number of applications including: information storage (advertising, museum art description), redirection to web sites, track and trace (for transportation tickets or brands), identification (flight passenger information, supermarket products) etc.

The popularity of these codes is mainly due to the following features: they are robust to the copying process, easy to read by any device and any user, they have a high encoding capacity enhanced by error correction facilities, they have a small size and are robust to geometrical distortions. However, those undeniable advantages also have their counterparts:

1) Information encoded in a QR code is always accessible to everyone, even if it is ciphered and therefore is only legible to authorized users (the difference between "see" and "understand").
2) It is impossible to distinguish an originally printed QR code from its copy due to their insensitivity to the Print-and-Scan (P&S) process.

In this paper, we propose to overcome these shortcomings by enriching the standard QR code encoding capacity. This enrichment is obtained by replacing its black modules by specific textured patterns. Besides the gain of storage capacity, these patterns can be designed to be sensitive to distortions due to the P&S process. These patterns, that do not introduce disruption in the standard reading process, are always perceived as black modules by any QR code reader. Therefore, even when the private information is degraded or lost in the copy, the public information is always accessible for reading.

The proposed two level QR (2LQR) code contains of: a first level accessible for any standard QR code reader, therefore it keeps the strong characteristics of the QR code; and a second level that improves the capacities and characteristics of the initial QR code.

The information in the second level is encoded by using $q-$ary ($q \geq 2$) code with error correction capacities. This information is invisible to the standard QR code reader because it perceives the textured patterns as black modules. Therefore, the second level can be used for private message sharing. Additionally, thanks to textured pattern sensitivity to P&S distortions, the second level can be used to distinguish the original 2LQR code from its copies.

The paper is organized as follows. We start with an introduction of QR code features and existing rich graphical codes in Section II. In addition, the distortion added during the P&S process will be discussed there. The proposed two level QR (2LQR) code as well as the proposed recognition method are presented in Section III. In Section IV, the experimental results show the efficiency of the proposed recognition methods and analyze the capacities of the proposed 2LQR code. Finally, Section V represents conclusions and perspectives.
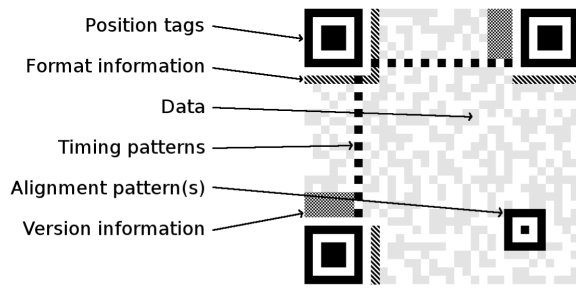
Fig. 1.   Specific QR code structure consists of position tags, alignment patterns, timing patterns, format information and version information patterns.

II. RELATED WORK

This section is split into three sub-sections (II-A, II-B, II-C). We start with a description of the standard QR code features in Section II-A. Definitions and descriptions of interesting rich graphical codes are presented in Section II-B. Finally, to highlight authentication of physical documents, the distortions added to any image during the P&S process are discussed in Section II-C.

A. QR Code Features

The QR code was invented for the Japanese automotive industry by Denso Wave[1] corporation in 1994. The most important characteristics of this code are small printout size and high speed reading process. The certification of QR code was performed by International Organization of Standardization (ISO), and its whole specification can be found in [3].

A QR code encodes the information into binary form. Each information bit is represented by a black or a white module. The Reed-Solomon error correction code [15] is used for data encryption. Therefore, one of 4 error correction levels has to be chosen during QR code generation. The lowest level can restore nearly 7% of damaged information, the highest level can restore nearly 30%.

Today, 40 QR code versions are available with different storage capacities. The smallest QR code version (version V1) has a 21×21 module size. It can store 152 bits of raw data at the lowest correction level. The biggest QR code version (version V40) has a 177×177 module size. It can store a maximum of 7089 bits of raw data at its lowest correction level.

As illustrated in Fig. 1, the QR code has a specific structure for geometrical correction and high speed decoding. Three position tags are used for QR code detection and orientation correction. One or more alignment patterns are used to code deformation adjustment. The module coordinates are set by timing patterns. Furthermore, the format information areas contain error correction level and mask pattern. The code version and error correction bits are stored in the version information areas.

The QR code generation algorithm consists of information encoding using Reed-Solomon error correction code, information division on codewords, application of mask pattern, placement of codewords and function patterns into the QR code. The QR code recognition algorithm includes the scanning process, image binarization, geometrical correction and decoding algorithm.

B. Rich Graphical Codes

In order to improve the graphical code properties, several rich graphical codes have recently been introduced. These rich graphical codes aim to add visual significance, to personalize the stored information or to increase the storage capacities. In this section, the different kinds of rich QR codes and several interesting rich graphical codes are presented.

The most simple type of rich QR codes is the user-friendly QR code. The target of these codes is to improve the aesthetic view of QR codes. It consists of changing the colors and shape of the modules, or of adding an image into the QR code. Different design QR code generators are proposed as free or paid applications.[2] However, most of these generators prefer to sacrifice the possibility of error correction for attractive design. Recently, the rich QR code, which adds the significance without losing error correction capacity, was introduced in [4]. The authors proposed a novel method of blending a color image into the QR code, which modifies the QR code source pixels so that the white (rsp. black) module pixels are transformed from white (rsp. black) to any RGB values and whose luminance value is considered as white (rsp. black) pixel by QR code binarization method.

Another way to personalize the stored information has been proposed under the name of contextual QR code [14]. It relates to the static QR code information with a particular context. The authors developed a specific application, which takes into account the individual users parameters (time, device type, IP address, location) in order to personalize (add the name of a user, change the language) an output message and to transmit user information into a server database.

The most popular enrichment is the stored capacity improvement. The HCC2D code [12] is a rich QR code which significantly increases the storage capacity of the standard QR code. The authors increased the density and storage capacity of standard QR code by replacing binary colored modules by RGB colored modules. The HCC2D code encodes information using 4, 8 or 16 module colors. This code inherits all the strong properties of standard QR codes, but it is not readable by a standard QR code reading application and needs to be printed using a color printer. One of the application scenarios for HCC2D code is facial biometrics [13].

Recently, the QR code steganography, which aims to hide a secret message into a QR code, was introduced. In [6] and [10] the authors suggest to insert the secret message by using the error correction capacity of the QR code. That means, they changed the bits encoded in the standard QR code, and inserted errors into it. In this case, the secret message does not disturb the reading process of the QR code message, but the error capacity of QR code is low. The maximum secret message length, mentioned in [6], is equal to 1215 bytes for QR code $V40$. There are also some approaches that embed an invisible watermark into the QR code image: the authors use

---

[1] http://www.qrcode.com/en/index.html

[2] An example can be found here: https://www.unitag.io/qrcode

the discrete cosine transform in [23] and the discrete wavelet transform in [17].

Let us take a look at some other interesting rich graphical codes. The multilevel 2D barcode [21] significantly improves the storage capacity of 2D code. The authors proposed to use one halftone cell to represent one multilevel 2D symbol. Only the black-and-white (B&W) halftone printers and low-resolution CDD-based scanners (up to 600 dpi) were used for their experiments. It was discovered that the rate of encoded version of multilevel 2D barcode is approximately 261 $byte/cm^2$ at a bit error rate of 4×10^{-2} in comparison with the rate of DataMatrix, which approximately equals to 58 $byte/cm^2$.

The rich DataMatrix code, named unsynchronized 4D barcode [8], increases the storage capacity by using RGB colors of modules and time. It consists of the 9 colored DataMatrix codes displayed in sequence on the screen. The particular color reference surrounds the DataMatrix and changes for each of 9 DataMatrix codes. This code has good storage capacity improvements, but it cannot be printed. The storage capacity is a transmission capacity and achieves 1400 characters per minute (23 characters per second) the success rate equals to 82%. This code is not readable by a standard DataMatrix reading application.

The authors in [7] proposed to use the graphical codes for authentication. The graphical code used is the copy detection pattern [11], which is a maximum entropy image, generated using a secret key, password or random seed. The authentication process is performed by the comparison of an original graphical code with the P&S graphical code embedded in the document. If the difference among these images is less than threshold λ, then this graphical code and document are authentic.

To summarize, we state that the length of a secret message, hidden in existing rich QR codes, is quite limited. In addition, there are no rich QR codes, that are sensitive to the copy process and can be used for whole document authentication. And finally, a rich graphical code, that has both properties, has never been suggested before.

C. P&S Process Impact

Any QR code production implies a printing process and a scanning process. The P&S process in authentication scenarios are considered as a physical unclonable function [7]. The textured patterns, that we propose to use in 2LQR code, are sensitive to the P&S process. In this section, the changes added during the P&S process are discussed, in order to understand why the output images suffer from this process.

The P&S process produces visible and invisible image modifications, which can be caused by resampling inherent to the P&S process, inhomogeneous lighting conditions, ink dispersion, varying speeds of the scanning device [5]. The modifications provided by the printer are not separable from modifications added by the scanner, that is why the distortions belong to both of them [24].

The most important elements of the *printing process* are printer resolution, digital halftoning, toner distribution,

physical construction and type of paper. The printer resolution is specified in dots per inch (dpi) and it influences the quality of the printed image. The digital halftoning is a transfer of gray-scale image to black-and-white image. This operation adds quantization noise in the image high frequencies [16] and introduces the image low pass filtering [19]. The non-uniform distribution of toner is the manufacturer defect and introduces geometrical image displacement [24]. The printing quality depends on the type of paper. Coated papers reduce dot gain by restricting ink absorption into the surface of the paper. Uncoated papers do not have additional layer, therefore the inks dry by being absorbed into the paper.[3] But as it was proved practically in [16], the printing operation in itself does not cause blurring, but it adds correlated noise, which varies every time, when something is printed.

The *scanning process* is specified by scanner resolution, gamma correction and scanner optics. The optical modulation transfer function of the scanner determines the scanner resolution (which is defined by the number of scanned pixels per inch) and is modeled as a Gaussian blur [24]. Therefore, the scanning process blurs the image. To correctly display the scanned image on the monitor, the gamma correction is applied to the image data generated by the scanner. This means that the data is raised to a power $1\backslash\gamma$, where γ is the parameter of a power function, which represents an intensity to voltage response curve. This process introduces non-linear transformation [16]. The CCD scanner sensors introduce thermal and dark noise, and the stepped motion jitter of the carriage introduces a Gaussian random noise [24].

All these changes mentioned produce pixel value distortions. These kinds of distortions were modeled in [9] and in [22]. Nevertheless, the P&S process also introduces geometric boundary distortion. This distortion is caused by rotation, scaling and cropping. This operation depends on the user and changes constantly.

This short discussion shows the rich graphical codes topicality, research interest and the variety of application scenarios.

III. RICH QR CODE WITH PUBLIC AND PRIVATE LEVELS

A rich QR code with a public level and a private level is discussed in this section. Two application scenarios for 2LQR code can be suggested: a private message sharing scenario and an authentication scenario. The main purpose of a private message sharing scenario is the invisible storage and transmission of private information into QR code. In a printed document authentication scenario, we aim to verify whether the printed document is an original or a copy. Only the original document (printed by authorities) is considered as an authentic. These application scenarios, as well as, the 2LQR code generation steps are described in Section III-A. Afterwards, we talk about 2LQR code capacities in Section III-B. To finish, the pattern recognition method is introduced in Section III-C.

A. Two Level QR (2LQR) Code Generation

Section II-A describes the specific structure of the standard QR code. Like the standard QR code, the 2LQR code has the

---

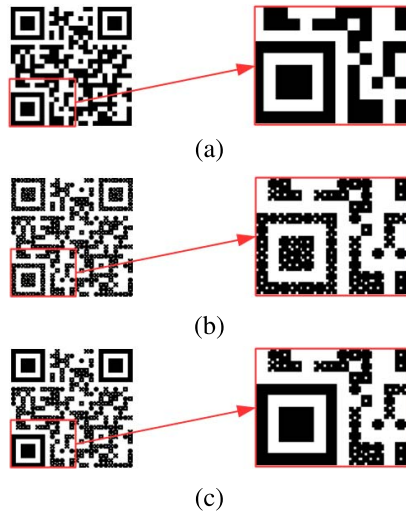[3]https://www.summitprintingpro.com/resources/paper-types.html

Fig. 2. A comparison of a) Standard QR code, with b) Proposed 2LQR code for private message sharing scenario and c) Proposed 2LQR code for authentication scenario.

Fig. 3. Overview of 2LQR code generation steps.

same specific structure, which consists of position tags, alignment patterns, timing patterns, version and format patterns. However, in the standard QR code, we have white and black modules and in the 2LQR code we have white modules and textured modules instead of black modules. This replacement of black modules by textured modules does not disrupt the standard QR code reading process. But it allows us to have a second storage level, which is invisible to the standard QR code reader. This second level contains the private message, encoded with q-ary ($q \geq 2$) code with error correction capacity. The textured modules are named textured patterns in the rest of this paper. These textured patterns have specific features and are used for private message M_{priv} storage in the proposed 2LQR code.

We suggest to use the 2LQR code for two scenarios: for private message sharing and for document authentication. The structure of the code in both scenarios is the same with slight differences on position tags. In private message sharing scenario, the black modules of these patterns are also replaced by textured patterns. However, in document authentication scenario, the position tags are not changed (i.e. they remain black modules).

The comparison of standard QR code with the proposed 2LQR code is illustrated in Fig. 2. The standard QR code is illustrated in Fig. 2.a. Likewise the proposed 2LQR code for private message sharing and for authentication are illustrated in Fig. 2.b and Fig. 2.c. The replacement of black modules in the standard QR code by textured patterns in 2LQR codes is shown. In both 2LQR codes, the alphabet dimension is equal to $q = 3$ (i.e. 3 different textured patterns are used for private level generation). Nevertheless, the alphabet can be extended by rising the pattern number used.

In the *private message sharing scenario*, the samples of used textured patterns are stored in the position tags. These textured patterns are mixed by using permutation σ and then are placed in the position tags. These textured patterns will be used for pattern detection. In this case, the reading algorithm can
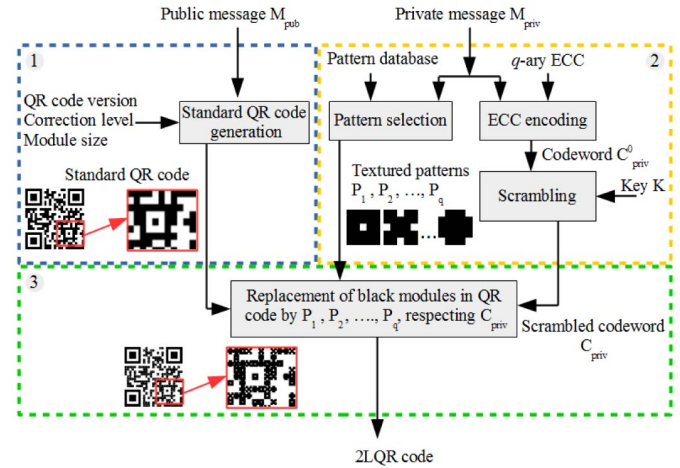
decode the private message M_{priv}, if the permutation σ and the ECC algorithm used for message encoding are known. For this, the proposed recognition method has to reconstruct the classes of textured patterns by using permutation σ, then create the characterization patterns for each used textured pattern, and compare the characterization patterns with textured patterns used in 2LQR code in order to recognize the patterns.

Unlike the previous scenario, in *document authentication*, it is necessary to verify the authenticity of the textured patterns used. Therefore, the textured patterns used in 2LQR code have to be compared with the original patterns. That is why, the black modules in the position tags do not changed.

Now we describe the 2LQR code generation steps as illustrated in Fig. 3. The input information is the public message M_{pub} and the private message M_{priv}. The output is the 2LQR code.

Public message M_{pub} storage (Fig. 3, block 1). The public message M_{pub} is stored in the standard QR code, using the classical generation method described in [3]. The standard QR code generation algorithm includes the following steps. First of all, the most optimal mode (numeric, alphanumeric, byte or Kanji) is selected by analyzing the message content. The message M_{pub} is encoded using the shortest possible string of bits. This string of bits is split up into 8 bit long data codewords. Then, the choice of error correction level is performed and the error correction codewords using the Reed-Solomon code are generated. After that, the data and error correction codewords are arranged in the correct order. In order to be sure that the generated QR code can be read correctly, the best (for encoded data) mask pattern is applied. After this manipulation, the codewords are placed in a matrix in a zigzag pattern, starting from the bottom-right corner. The final step is to add the function patterns (position tags, alignment, timing, format and version patterns) into the QR code.

Private message M_{priv} encoding (Fig. 3, block 2). The private row-bit string is encoded using error correction code (ECC) to ensure the message error correction after the P&S operation. We use the block codes, and more precisely cyclic codes (or polynomial-generated codes) such as Golay code [15] or Reed-Solomon code, for message encoding.

Cyclic codes can be defined in matrix form and polynomial form. Any cyclic code C is defined by $[n, k, d]$ parameters, where n is the length of the codeword, k is the number of information digits in the codeword, d is the minimum distance between distinct codewords. The $n - k$ digits in the codeword are called parity-check digits, and in ECC these digits are used for error detection and correction. The minimum distance d of code C ensures that up to $t = (d - 1)/2$ errors can be corrected by the code C.

Let $R = A[x]/(x^n - 1)$ be a polynomial ring over a Galois field $A = GF(q)$. The cyclic code C elements are defined with polynomials in R so that the codeword $(c_0, c_1, \ldots, c_{n-1})$ maps to the polynomial $c_0 + c_1 x + \cdots + c_{n-1} x^{n-1}$, and the multiplication by x corresponds to a cyclic shift. The code C is generated by a generator polynomial $g(x)$, which is the code polynomial of the minimum degree in a (n, k) cyclic code C. Therefore, the generator polynomial $g(x)$ is a factor of polynomial $x^n - 1$.

Let k informative digits of message be represented by a polynomial $m(x)$, of degree, at most $k - 1$. Then the codeword $c(x)$ is the polynomial of the form:

$$c(x) = m(x)g(x),$$
$$c(x) = c_0 + c_1 x + \cdots + c_{n-1} x^{n-1}.$$

Therefore, the encoded informative digits are $(c_0, c_1, \ldots, c_{n-1})$.

The private message $M_{priv} = (m_1^{priv}, \ldots, m_k^{priv})$ is encoded using ECC $[n, k]$. First of all the M_{priv} is represented in polynomial form $m_{priv}(x)$. Then, the polynomial form of the codeword $c_{priv}(x) = m_{priv}(x)g(x)$ is calculated that represents the codeword C_{priv}^0.

After that, as illustrated in Fig. 3, the codeword C_{priv}^0 is scrambled using the key K. Thus, the scrambled codeword digits are:

$$C_{priv} = (c_0^{priv}, c_1^{priv}, \ldots, c_{n-1}^{priv}).$$

Textured Pattern Selection: The textured patterns $P_i, i = 1, \ldots, q$ are images of size $p \times p$ pixels. We choose q patterns from a database of $Q \gg q$ textured patterns, which are binary and have the same density (ratio of black pixels), equal to b, and have related spectra. The reading capacity of private level depends on pattern density: a large density value can disable the reading process of private level.

Let $S_i, i = 1, \ldots, q$ be the P&S degraded versions of textured patterns $P_i, i = 1, \ldots, q$. The Pearson correlation between a pattern P_i and a P&S pattern S_i is:

$$cor(P_i, S_i) = \frac{\sum_w \sum_h (P_i^*(w, h))(S_i^*(w, h))}{\sqrt{\sum_w \sum_h (P_i^*(w, h))^2} \sqrt{\sum_w \sum_h (S_i^*(w, h))^2}},$$
(1)

where $P_i^*(w, h)$ (rsp. $S_i^*(w, h)$) are the central values of P_i (rsp. S_i) defined by $P_i^*(w, h) = P_i(w, h) - \mu_{P_i}$ (rsp. $S_i^*(w, h) = S_i(w, h) - \mu_{S_i}$) with $\mu_{P_i} = \frac{1}{p^2} \sum_w \sum_h P_i(w, h)$ (rsp. $\mu_{S_i} = \frac{1}{p^2} \sum_w \sum_h S_i(w, h)$).

The same textured patterns were used for the generation of the textured image in [18]. That is why only the patterns

which respect the two following conditions [18], could be used in 2LQR code generation:

1) Each textured pattern $P_i, i = 1, \ldots, q$ has to be better correlated with its P&S degraded version S_i, $i = 1, \ldots, q$ than with all other P&S degraded versions $S_j, j = 1, \ldots, q, i \neq j$:

$$\forall i, j \in \{1, \ldots, q\}, i \neq j, cor(P_i, S_i) > cor(P_i, S_j).$$
(2)

2) The P&S degraded version $S_i, i = 1 \ldots q$ of each pattern has to be better correlated with its original pattern $P_i, i = 1 \ldots q$ than with all other original patterns S_j, $j = 1 \ldots q, i \neq j$:

$$\forall i, j \in \{1, \ldots, q\}, i \neq j, cor(P_i, S_i) > cor(P_j, S_i).$$
(3)

These two conditions (2)−(3) can be rewritten in the form:

$$\forall i \in \{1, \ldots, q\}, cor(P_i, S_i) = \max_{\forall j \in \{1, \ldots, q\}} (cor(P_i, S_j))$$
$$= \max_{\forall j \in \{1, \ldots, q\}} (cor(P_j, S_i)). \quad (4)$$

The condition (4) is valid for all values of j ($\forall j \in \{1, \ldots, q\}$), but if the value $j = i$ is excluded, the condition (4) can be rewritten in the form:

$$\forall i, j \in \{1, \ldots, q\}, cor(P_i, S_i) = \max_{i \neq j}(cor(P_i, S_j)) + \epsilon_{i1}$$
$$= \max_{i \neq j}(cor(P_j, S_i)) + \epsilon_{i2}. \quad (5)$$

The condition (5) was proposed due to the recognition method, which is based on maximising the correlation values between the original patterns and its P&S degraded versions. Therefore a new criteria which is the minimum distance between the best correlation score and the second best one can be added. This distance should be greater than a given ϵ threshold. That is why the limits for distances ϵ_{i1} and ϵ_{i2} are set:

$$\epsilon_{i1} \geq \epsilon \quad \text{and} \quad \epsilon_{i2} \geq \epsilon, \quad (6)$$

the theoretical distance values have to be in the interval $\epsilon_{i1}, \epsilon_{i2} \in [\epsilon, 2]$.

Therefore, only the textured patterns, which respect the conditions (4)−(6), can be combined, and used for 2LQR code generation.

Black module replacement (Fig. 3, block 3). The codeword C_{priv} is inserted in standard QR code by replacing the black modules with textured patterns $P_1, \ldots, P_q$ respecting the codeword C_{priv}, starting from the bottom-right corner. Then, in the case of *private message sharing scenario*, the textured patterns are placed in the position tags with respect to the chosen permutation σ, see Fig. 2.b. In the case of *authentication scenario*, the standard position tags keep unchanged black modules, see Fig. 2.c.
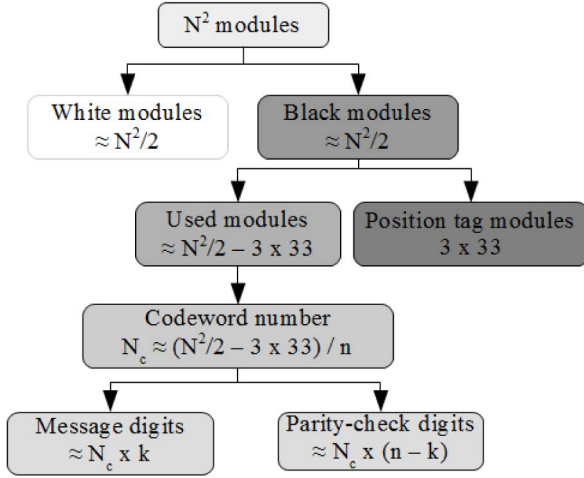
Fig. 4. Storage capacity of the 2LQR code for QR code of size $N \times N$ modules, where n is the codeword length, k is the number of information digits, $n - k$ is the number of parity-check digits.

B. Storage Capacity of 2LQR Code

In this section the storage capacities of proposed 2LQR code illustrated in Fig. 4 is discussed. Let N^2 be the number of modules in a standard QR code. As QR code construction aims to have an approximately equal number of black and white modules, we can suppose that $N^2/2$ is approximately the number of black modules in standard QR code. It has three position tags, each tag has 33 black modules. That is why, there are approximately $(N^2/2 - 3 \times 33)$ black modules in QR code, that could be replaced by textured patterns.

Let n be the total number of digits in a codeword, k be the number of message digits, $n - k$ be the number of error correction bits in a codeword. Therefore, the number of codewords, that could be inserted in the second level of a 2LQR code, is approximately equal to $N_c \approx (N^2/2 - 3 \times 33)/n$. And the number of message digits is approximately equal to $N_c \times k$, that is why the length of message is approximately equal to $log_2(q) \times N_c \times k$, where q is the alphabet dimension.

These formulas allow us to calculate the maximal storage capacity of the 2LQR code for a fixed module number N^2 i.e. for a fixed QR code version. At the same time, these formulas define the version of the QR code, which can be used for insertion into a fixed message digit number.

Example: Let take the Golay binary code [23, 12, 7], i.e. $k = 12, n - k = 11, n = 23$, and the standard QR code version $V2 : N^2 = 25 \times 25 = 625$. The number of codewords is approximately equal to $N_c = (625/2 - 99)/23 \approx 9$. Therefore, we can insert $N_c \times k = 9 \times 12 = 108$ binary digits (bits) of private message.   □

C. Recognition Method

The QR code reproduction implies the use of the printing process and the scanning process. As it was discussed in Section II-C, the P&S process is a complicated process, which blurs and modifies the output image. We recall that S_i, $i = 1, \ldots, q$ is the P&S degraded version of the textured pattern $P_i, i = 1, \ldots, q$.
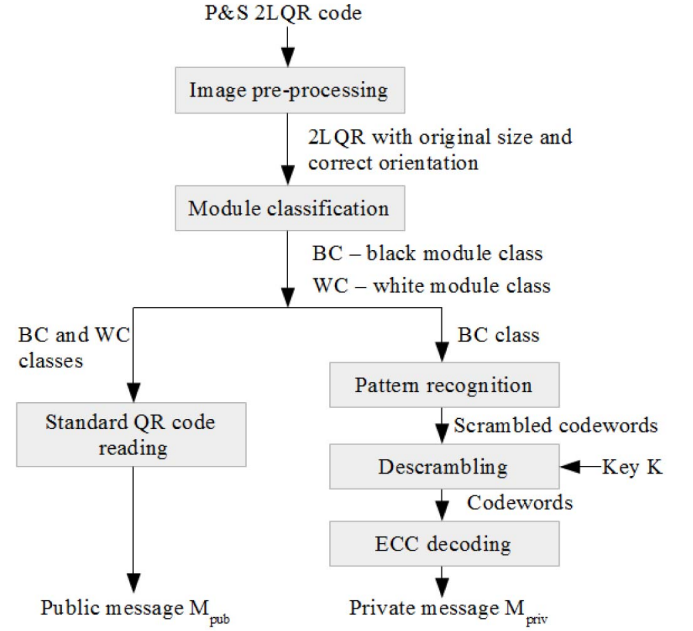

Fig. 5. The overview of 2LQR code reading process.

The overview of the 2LQR code reading process is illustrated in Fig. 5. First, the geometrical distortion of P&S 2LQR code has to be corrected during the **pre-processing step**. The position tags are localized by the standard process [3] to determine the position coordinates. The linear interpolation is applied in order to re-sample the P&S 2LQR code. Therefore, at the end of this step, the 2LQR code has the correct orientation and original size $N \times N$ pixels.

The second step is the **module classification** performed by any threshold method. We use global threshold, which is calculated as a mean value of the whole P&S 2LQR code. Then, if the mean value of the block $p \times p$ pixels is smaller than global threshold, this block is in a black class (BC). Otherwise, this block is in a white class (WC). The result of this step is two classes of modules.

In the next step, two parallel procedures are completed. On one side, **the decoding of public message** M_{pub} is performed by using standard QR code decoding algorithm [3] and the positions of the white and black modules. And on the other side, the BC class is used for **pattern recognition** of the textured pattern in P&S 2LQR code. The class BC contains the textured patterns $BP_i, i = 1, \ldots, N_c \times n$, where $N_c \times n$ is the total number of codeword digits, N_c is the number of codewords, n is the number of digits in the codeword. Therefore, there are $N_c \times n$ textured patterns, which belong to q classes.

The proposed pattern detection method compares the P&S patterns ($BP_i, i = 1, \ldots, N_c \times n$) with characterization patterns by using the Pearson correlation (1). Since there are two different application scenarios, two types of characterization patterns have to be used.

In *private message sharing scenario*, the chosen patterns $P_1, \ldots, P_q$ are stored in position tags. We use these patterns in order to create the characterization patterns (see Section III-A and Fig. 2.b). The permutation σ has been used to mix the
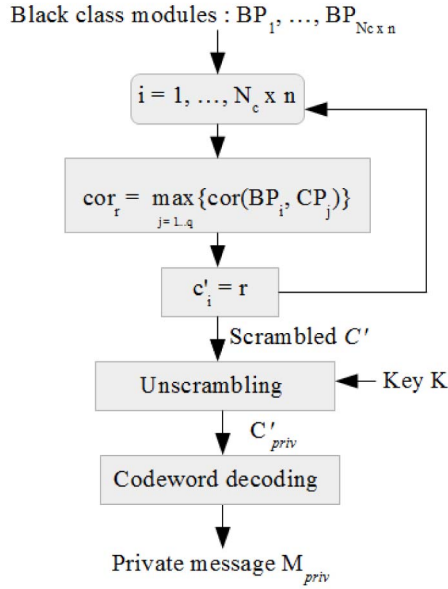
Fig. 6. The pattern recognition method scheme by using characterization patterns $CP_1, \ldots, CP_q$.

TABLE I

STORAGE CAPACITY INFORMATION OF PUBLIC AND PRIVATE LEVELS OF THE 2LQR CODE

	Public level	Private level
	(Standard QR code V2 Low)	(2LQR code V2 with Golay [11, 6, 5])
Modules	625	312
Encoded bits	324	216
Ternary message digits	–	114
Message bits	272	180

IV. EXPERIMENTAL RESULTS

This section illustrates both the generation steps of the 2LQR code in Section IV-A and the message extraction steps in Section IV-B. Then, the storage capacities of the 2LQR code is discussed in Section IV-C. After, the reading capacities of public and private levels are presented in Section IV-D. In addition, we argue the authentication properties of the proposed 2LQR code in Section IV-E.

Application Scenario: For example, the public level can store the Surname, First name, Date of Birth and Place of Birth of a person. Then, the secret information, which is the number of their bank account, encoded in the private level.

Setup: In these experiments, the version V2 of the QR code in Low error correction level is used. This version has 25×25 module size and can store 272 bits of a message.

Error Correction Code: The private information is encoded with the ternary Golay code [11, 6, 5]. Each codeword has $n = 11$ digit length, where $k = 6$ digits correspond to the message and $n - k = 5$ digits are parity-check digits. In the second level of version V2 of QR, there are approximately 216 black modules (if we do not use the black modules of position tags to store the information). Therefore, with the ternary Golay code [11, 6, 5] we can store 114 ternary digits, that corresponds to nearly 180 message bits. The storage capacities of public and private levels are determined in Table I.

A. 2LQR Code Generation

The 2LQR code generation, as was previously mentioned in Section III-A, consists of four steps: standard QR code generation, codeword generation, pattern selection and replacement of black modules in QR code.

Standard QR Code Generation: The standard QR code with public message M_{pub} "John Doe - 13/05/1958 - New York" is generated by using a free online QR code generator. The generated standard QR code, version $V2$, is illustrated in Fig. 7. The actual size of this QR code is $1.2 \times 1.2 \ cm^2$.

Codeword Generation: The private message M_{priv} with a 180 bit length, which corresponds to 114 ternary digits, is defined. We encode this message with ternary Golay code [11, 6, 5]. The codeword C^0_{priv} has 209 digit length. The last 7 accessible digits are defined randomly. Then, we use the key K for scrambling of the codeword C^0_{priv}, and the codeword C_{priv} is obtained.

Pattern Selection: We chose three textured patterns P_1, P_2 and P_3, which are illustrated in Fig. 8. These patterns have some particular characteristics:

- they have a size of 12×12 pixels;
- they are binary;

textured patterns in position tags. Now, we apply the inverse permutation σ^{-1} and obtain q sets (as q is the dimension of our alphabet) of the P&S representative patterns. Each position tag consists of 33 black modules (for any QR code version), in total there are 99 representative patterns. That is why the dimension of each representative set is equal to an integer $\upsilon = [99/q]$. We calculate the characterization patterns as:

- the *mean images* obtained by averaging the υ P&S textured patterns : $CP_l = \frac{1}{\upsilon} \sum_{\tau=1}^{\upsilon} S_l^\tau, l = 1, \cdots, q$;
- the *median images* obtained as median images of υ P&S textured patterns: $CP_l = median(S_l^1 \cdots S_l^\upsilon)$, $l = 1, \cdots, q$.

In the *authentication scenario*, since the black modules in position tags have not been replaced by textured patterns (see Section III-A and Fig. 2.c). We compare the P&S patterns with the original patterns, that is why the characterization patterns are replaced by:

- the *original images*: $CP_l = P_l, l = 1, \cdots, q$.

The flowchart of the proposed pattern recognition method is presented in Fig. 6. The class BC contains at least $N_c \times n$ textured patterns $BP_i, i = 1, \ldots, N_c \times n$. For each textured pattern $BP_i, i = 1, \ldots, N_c \times n$, a maximum correlation value cor_r with characterization patterns $CP_1, \ldots, CP_q$, i.e. $cor_r = \max_{j=1,\ldots,q} \{cor(BP_i, CP_j)\}$ is determined. Then, the codeword digit is equal to r: $c'_i = r$. So, the output of this method is the codeword $C'_{priv} = (c'_1, \ldots, c'_{N_c \times n})$.

The last steps of the 2LQR code reading process are **unscrambling** using key K and **ECC decoding** of the obtained codeword C'_{priv}. We use the parity-check digits for error detection and correction. For error correction and decoding, one of the classical ECC decoding algorithms (i.e. error syndrome decoding, maximum likelihood decoding algorithms) can be used. The result of this algorithm is the restored private message M_{priv}.

Fig. 7.   The example of a) Standard QR code with public message M_{pub}, b) Standard QR code at actual size defined at 600 dpi, $(1.2 \times 1.2 \ cm^2)$.
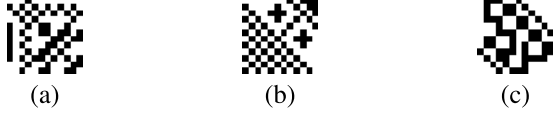

Fig. 8.   The three textured patterns used for private level generation: a) Pattern 1, b) Pattern 2, c) Pattern 3.

TABLE II

THE CORRELATION VALUES FOR THE ORIGINAL PATTERNS P_1, P_2, P_3 AND ITS P&S DEGRADED VERSIONS S_1, S_2, S_3, THE MINIMAL DISTANCES AMONG DIAGONAL CORRELATION VALUES AND ALL OTHER CORRELATION VALUES IN THE SAME LINE AND THE SAME COLUMN $\min\{\epsilon_{i1}, \epsilon_{i2}\}, i = \{1, \ldots, 3\}$

	P_1	P_2	P_3
S_1	**0.4432**	−0.0072	0.1016
S_2	−0.0054	**0.2754**	−0.0076
S_3	0.0293	−0.0860	**0.4752**
$\min\{\epsilon_{i1}, \epsilon_{i2}\}$	0.3416	0.2830	0.3737

- they have the same number of black pixels $b = 60$ (that corresponds to nearly 42%);
- they have spectra related between them.

The original and the P&S degraded versions of these patterns satisfy conditions (4)−(6) with $\epsilon = 0.25$. The correlation values, as well as, ϵ values are illustrated in Table II. The diagonal correlation values $cor(P_i, S_i), i = 1, \ldots, 3$, which are equal to $cor(P_1, S_1) = 0.4432$, $cor(P_2, S_2) = 0.2754$ and $cor(P_3, S_3) = 0.4752$, are the maximum in line and in column. That is why condition (4) is respected. Then, we calculate the distances among $cor(P_i, S_i)$ and $cor(P_i, S_j), i, j \in \{1, \ldots, 3\}, j \neq i$, i.e. $\epsilon_{i1} = cor(P_i, S_i) - cor(P_i, S_j)$ and $\epsilon_{i2} = cor(P_i, S_i) - cor(P_j, S_i), i, j \in \{1, \ldots, 3\}, j \neq i$. We notice that $\min\{\epsilon_{i1}, \epsilon_{i2}\} \geq \epsilon = 0.25$, as

$$\min\{\epsilon_{11}, \epsilon_{12}\} = 0.3416 \geq 0.25 = \epsilon,$$
$$\min\{\epsilon_{21}, \epsilon_{22}\} = 0.2830 \geq 0.25 = \epsilon,$$
$$\min\{\epsilon_{31}, \epsilon_{32}\} = 0.3737 \geq 0.25 = \epsilon.$$

That is why, the condition (6) is also respected. Therefore, these patterns can be used for second level generation.

The Replacement of Black Modules in the QR Code: The private level of the 2LQR code is constructed by the replacement of black modules by textured patterns illustrated in Fig. 8.a-c with respect to the codeword C_{priv}. We replace the black modules by columns starting from the bottom-right corner of the QR code.
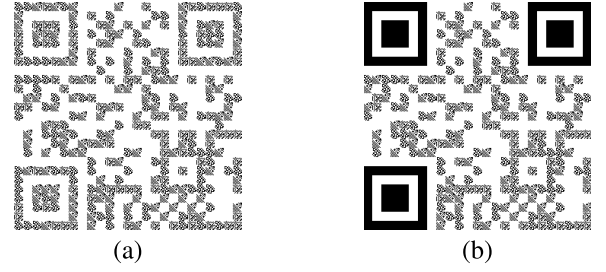

Fig. 9.   The example of a 2LQR code a) for private message sharing, b) for authentication. The alphabet dimension is equal to $q = 3$.
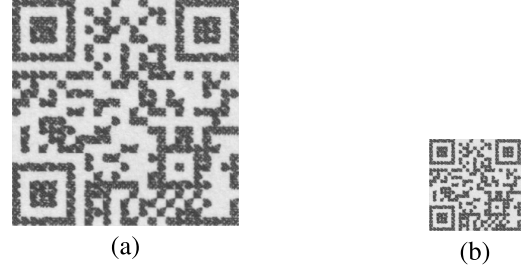

Fig. 10.   The example of a) P&S 2LQR code for private message sharing and b) P&S 2LQR code at actual size defined at 600 dpi.

In the *private message sharing scenario*, the black modules in position tags are replaced by patterns P_1, P_2 and P_3 with respect to permutation σ. The patterns stored in position tags will be used for the creation of characterization patterns $CP_i, i = 1, .., 3$. In the *authentication scenario*, the black modules of position tags are not changed. The example of the 2LQR code suitable for the *private message sharing scenario* is illustrated in Fig. 9.a, the 2LQR code suitable for the *authentication scenario* is illustrated in Fig. 9.b. The 2LQR code size is equal to $1.2 \times 1.2 \ cm^2$. The public level of this 2LQR code is readable by any QR code reader.

B. Message Extraction

For the pattern detection experiments, we printed the same 2LQR code 1000 times in 600 dpi using Brother HL-4150 printer. Then, we scanned each printed 2LQR code in 600 dpi using Canon LIDE 210 scanner. Fig. 10 illustrates an example of the P&S 2LQR code. In comparison with the original 2LQR code (Fig. 9), these images (Fig. 10.a-b) are blurred and in gray-level (instead of being binary).

For each P&S 2LQR code, the proposed detection method is applied with characterization patterns (mean and median) for the message sharing scenario and with the original patterns for the authentication scenario. The detection results are presented in Table III. The error probability of pattern detection by using characterization patterns is $0.06 - 0.12\%$ for message sharing and by using original patterns for authentication it is equal to 0.04%. In addition, we note that all the errors of pattern detection are caused by incorrect detection of the patterns P_2 and P_3. Therefore, the degraded version of these patterns are sometimes better correlated with two other patterns.

TABLE III

PATTERN DETECTION RESULTS AFTER P&S PROCESS IN THE 2LQR CODE WITH $q = 3$

% of patterns detected incorrectly	Message sharing Use of characterization patterns		Authentication
	Mean	Median	Original
% of P_1	0.00%	0.00%	0.00%
% of P_2	0.01%	0.03%	0.00%
% of P_3	0.16%	0.31%	0.12%
Error probability of pattern detection	0.06%	0.12%	0.04%

TABLE IV

ERROR PROBABILITY OF MESSAGE DECODING AFTER GOLAY ERROR CORRECTION ALGORITHM FOR $q = 3$

	Message sharing Use of characterization patterns		Authentication
Error probability of digit decoding after ECC	Mean	Median	Original
	0.00%	0.03%	0.00%

TABLE V

THE STORAGE CAPACITY CHANGES IN RESPECT OF q ($q = 2, 3, 8$) ALPHABET IN THE 2LQR CODE

Alphabet dimension q	Message bit number in		Total number of message bits
	public level	private level	
$q = 2$	272	108	380
$q = 3$	272	180	452
$q = 8$	272	810	1082

We apply the unscrambling operation using key K to the sequence of numbers, which corresponds to detection patterns. Since our private message was encoded using ternary Golay ECC, the error correction and decoding algorithm is applied to get the private message M_{priv}. The error probabilities of incorrect digit decoding are presented in Table IV. We conclude that all errors caused by incorrect pattern detection are corrected by Golay error correction algorithm: the 100% of the message digits are decoded.

C. Storage Capacity Analysis

In this section, we aim to study the storage capacity of the 2LQR code, using a fixed surface equal to 1.2×1.2 cm^2 and a fixed pattern density equal to approximately 42%. The storage capacity of the 2LQR code can be increased both by increasing the value of q, which is the number of digits and textured patterns and by decreasing the size $p \times p$ pixels of the textured patterns. In both cases, some problems in pattern detection are identified. Therefore, we have to find a pattern size and a pattern number trade-off.

In the first experiment, we increase the storage capacity of the 2LQR code by increasing the dimension q of the alphabet set. The public message is stored in the version V2 of QR code. Therefore, the length of the public message M_{pub} is constant and equal to 272 bits. Then, we use binary Golay [23, 12] ($q = 2$), ternary Golay [11, 6] ($q = 3$) and $8-$ary Reed-Solomon [7, 3] ($q = 8$) error correction codes for the encoding of the private message M_{priv}. The number of message bits is increased by increasing the dimension q of the alphabet, see Table V. For example the use of binary alphabet $q = 2$ can increase the storage capacity from 272 bits up

TABLE VI

THE ERROR PROBABILITY OF PATTERN DETECTION AND DIGIT DECODING AFTER ECC ALGORITHM FOR DIFFERENT ALPHABET DIMENSIONS ($q = 2, 3, 8$)

Alphabet dimension q	Error probability of					
	pattern detection			digit decoding		
	Mean	Median	Original	Mean	Median	Original
$q = 2$	0.00%	0.00%	0.00%	0.00%		
$q = 3$	0.06%	0.08%	0.02%	0.00%		
$q = 8$	0.39%	0.51%	0.11%	0.00%		

TABLE VII

THE STORAGE CAPACITY CHANGES IN RESPECT OF PATTERN SIZE DECREASING IN 2LQR CODE WITH $q = 3$

Pattern size (pixels)	QR code		Message bit number		Total message bit number
	version	size	public	private	
12×12	$V2$	300×300	272	180	452
8×8	$V5$	296×296	864	504	1368
6×6	$V8$	294×294	1552	950	2502

to 380 bits, the use of ternary alphabet $q = 3$ increases the storage capacity up to 452 bits and the use of $8-$ary alphabet $q = 8$ increases the storage capacity up to 1082 bits.

We have 1000 P&S samples of each 2LQR code (i.e. the 2LQR codes with the number of textured patterns equal to $q = 2$, $q = 3$ and $q = 8$). The proposed pattern detection method and the ECC decoding algorithm are applied to each sample. Table VI presents the error probability of pattern detection and digit decoding results. We can see that the number of incorrect pattern detection increases when the alphabet dimension q increases, but these errors do not occur after the decoding of the second level. The error probability of pattern detection is equal to 0%, when the number of pattern used is $q = 2$. In the same time the increase of pattern number decreases the pattern detection results: the error probability of pattern detection is equal to $0.11 - 0.51\%$ for $q = 8$.

In the second experiment, we increase the storage capacity of the 2LQR code by decreasing the size of the textured patterns. The patterns in Section IV-A have a size of 12×12 pixels, and the real size of QR code is 300×300 pixels or 1.2×1.2 cm^2. In this experiment, we fix the size of the QR code approximately equal to 1.2×1.2 cm^2 and the alphabet dimension $q = 3$, but the size of patterns and the version of the QR code are varied. This comparison is performed using a pattern size of 12×12 pixels and QR code version V2, a pattern size of 8×8 pixels and QR code version V5, and a pattern size of 6×6 pixels and QR code version V8. This approach can significantly increase the storage capacity of the 2LQR code, as illustrated in Table VII. The storage capacity of QR code $V2$ is increased from 272 bits up to 452 bits by using textured patterns with a size of 12×12 pixels. The patterns with a size of 6×6 pixels increase the storage capacity of QR code $V8$ from 1552 bits to 2502 bits.

Similarly to the first experiment, we have 1000 P&S samples of each 2LQR code type (i.e. the 2LQR codes with textured patterns of size 6×6, 8×8 and 12×12 pixels). The proposed pattern detection method and the ECC decoding algorithm are applied for each sample. The comparison of the error probability of each pattern detection and digit decoding is presented in Table VIII. The decreasing of pattern size slightly

TABLE VIII

THE ERROR PROBABILITY OF INCORRECT BIT DECODING AFTER ECC ALGORITHM FOR DIFFERENT SIZES OF TEXTURED PATTERNS WITH $q = 3$

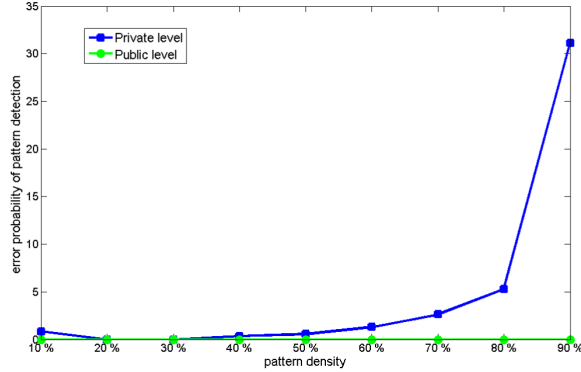| Pattern size (pixels) | Error probability of | | | | | |
| | pattern detection | | | digit decoding | | |
	Mean	Median	Original	Mean	Median	Original
12×12	0.06%	0.08%	0.02%	0.00%	0.00%	0.00%
8×8	3.14%	2.95%	0.23%	2.33%	2.31%	0.00%
6×6	6.45%	6.32%	1.06%	6.47%	5.98%	0.08%

Fig. 11. The error probability of module detection for public level (green lines) and of textured pattern detection for private level (blue line) depending on pattern density (from 10% to 90%).

decreases the pattern detection and the successful decoding of the message. Nevertheless, we can correctly decode the message encoded in the 2LQR code by using 8×8 size textured patterns. The error probability of the second level digit decoding by using 6×6 size patterns is equal to 0.08%, that means a small part of the private message is lost. The error probability of pattern detection varies from 0.02% until 6.45% depending on the size of the textured patterns.

D. Reading Capacities of First and Second Levels

In this section, we study the reading capacities of the first and second levels of the 2LQR code depending either on pattern density or on pattern size.

Firstly, we change the **pattern density** from 10% to 90%, but fix the pattern size to 8×8 pixels and the actual size of the 2LQR code to 1.2×1.2 cm^2. The pattern detection results are illustrated in Fig. 11. The public level (the information stored in the standard QR code) is readable all the time, even when the density is low. At the same time, the number of textured pattern detection errors increases when the density in the private level is increased. Nevertheless, thanks to error correction code, the correct private message can be extracted even for patterns with a density of 80%.

Secondly, we vary the **pattern size** from 3×3 pixels to 12×12 pixels, but fix the pattern density equal to approximately $40 - 42\%$ and the 2LQR code actual size equal to 1.2×1.2 cm^2. The pattern detection results are illustrated in Fig. 12. Here, the first level is readable for patterns that have a size from 4×4 pixels to 12×12 pixels. The second level is readable from a 5×5 pixel pattern size, using the error correction capacity.
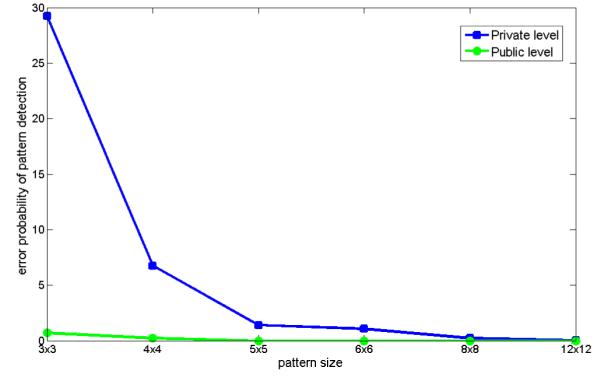

Fig. 12. The error probability of modules detection for public level (green lines) and of textured pattern detection for private level (blue line) depending on pattern size (from 3×3 pixels to 12×12 pixels).

TABLE IX

RICH GRAPHICAL BARCODE COMPARISON

| Code name | Storage capacity ($bits/inch^2$) | | | Color printing | Copy sensitivity |
	public	private	total		
HCC2D code [12]	15048	0	15048	Yes	No
Multilevel 2D barcode [21]	11224	0	11224	No	No
Graphical code for authentication [7]	0	0	0	No	Yes
QR code with hidden message[4] [6]	7548	3102	10650	No	No
Proposed 2LQR code[4]	7548	6386	13934	No	Yes

Experimental results show that for the correct reading process of public and private levels:
- The pattern size should be from 6×6 pixels to 12×12 pixels (the larger pattern size can significantly increase the QR code printing surface).
- The pattern density can vary from 20% to 80%. But we have to note that the image contrast is too weak, when the pattern density is less than 40%. And when the pattern density is larger than 70%, the distance among the correlation values is quite weak.
- The alphabet dimension q can be increased up to $q = 8$. The larger alphabet dimensions could disrupt the textured pattern detection due to the P&S impact.

Summarizing these experiments, the robustness of the proposed 2LQR code depends on the textured pattern size, the alphabet dimension q and the textured pattern density. For example, if we want to increase the storage capacity of the 2LQR code (either by decreasing the pattern size or increasing the alphabet dimension), the robustness (pattern recognition rate) will decrease (significantly or not). Analogically, the boundary values of textured pattern densities (significantly low or high) decrease the 2LQR code robustness. Therefore, we should take into account two trade-offs: storage capacity - robustness trade-off and QR code contrast - robustness trade-off.

To conclude the proposed 2LQR code evaluation, we compare it with several existing rich graphical barcodes in Table IX. Only two (including the proposed 2LQR code)

---

[4]Storage capacity calculated for QR code V40 fixing the barcode size to 3.1329 $inch^2$.
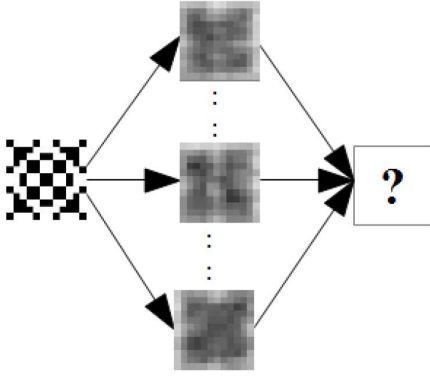
Fig. 13. Examples of modifications added to textured patterns during the P&S process.

of the mentioned graphical codes are sensitive to copying process, as well as this, only two have the private storage level capability. The maximum storage capacity of QR code with hidden message [6] is equal only to 9720 bits using QR code V40. The proposed 2LQR code increases the hidden message length up to 20000 bits when it is constructed using an 8-ary Reed-Solomon ECC in a QR code $V40$.

Therefore, the strongest main points of the proposed 2LQR code method in comparison with mentioned graphical code methods is the combination of four characteristics which are the two storage levels, the readability of first level by standard QR code application, the increased storage capacity and the sensitivity to copying process.

E. The Proposed 2LQR Code and Its Authentication

We have to differentiate document authentication from data (document content) authentication. The document hash function can authenticate data. That is why, recently, the document tamper proofing scenario has been proposed [20] for local document content authentication. The document is divided into parts and the local hash is computed for each part. Then, these hashes can be stored in 2D barcodes.

We suggest to use the proposed 2LQR code for document tamper proofing. In addition, due to the specific characteristics of the used textured patterns, the original 2LQR code can be distinguished from one of its copies to ensure document authentication. This functionality has been performed due to the impact of the P&S process, that can be considered as physically unclonable because of both the deficiencies of the physical process and the stochastic nature of the matter [7].

As we mentioned in Section II-C, any P&S process adds specific changes in each image. These modifications can be provided by ink dispersion (in the paper or onto the device output), non homogeneous luminosity conditions during the scanning process, inherent re-sampling of the P&S process or variable speed during the acquisition process [5]. An example of modifications which the textured patterns go through is illustrated in Fig. 13. Due to the P&S impact, it is difficult to model the P&S degraded versions of proposed textured patterns.

We decided to measure the difference between the document after a P&S process and a copy of the document after
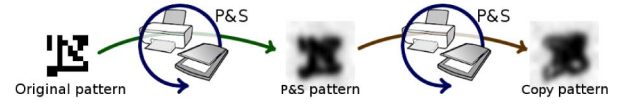

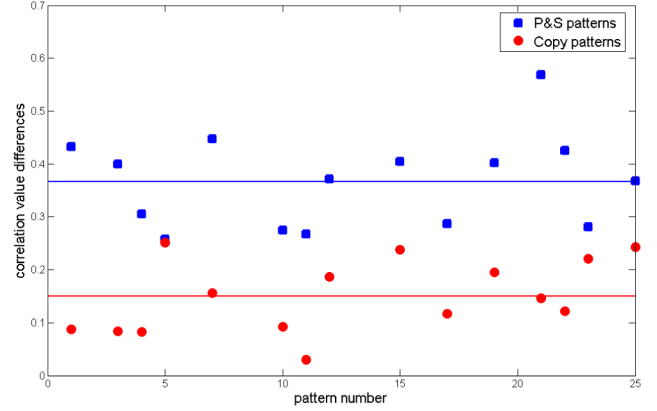Fig. 14. Production of P&S pattern and copy pattern from the original pattern.

Fig. 15. Degradation of textured patterns by the P&S process: the values ϵ are differences of correlation values with original patterns P_1, P_2 and P&S patterns S_1, S_2, the values ϵ_c are differences of correlation values with original patterns P_1, P_2 and copy patterns S_1^c, S_2^c. The blue (red) line is the mean value of differences ϵ (ϵ_c).

two P&S processes. A numeric original pattern is called an original pattern, a pattern after the P&S process is called a P&S pattern, and a pattern after two P&S processes is called a copy pattern, see Fig. 14.

For this experiment, we generate a 2LQR code with binary alphabet $q = 2$, i.e. only two textured patterns P_1, P_2 are used for private level generation. This 2LQR code was printed and scanned once (we obtain a P&S document) and then, this P&S document was again printed and scanned a second time (we obtained a copy document). Therefore, there were three documents: original document, P&S document and a copy of the P&S document. We compared the original patterns with the P&S patterns, and then, the original patterns with the copy patterns by using distances between correlation values ϵ:

$$\epsilon = |cor(P_1, S_i) - cor(P_2, S_i)|,$$
$$\epsilon_c = |cor(P_1, S_i^c) - cor(P_2, S_i^c)|,$$

where S_i is the P&S version of P_i, S_i^c is the copy version of P_i, $cor(P_1, S_i)$ (resp.$cor(P_2, S_i)$) is the Pearson correlation value between the original P_1 (P_2) textured patterns and the P&S version of textured pattern, $cor(P_1, S_i^c)$ (resp.$cor(P_2, S_i^c)$) is the Pearson correlation value between the original P_1 (P_2) textured patterns and the copy version of the textured pattern, $i = 1, 2$. These correlation value differences are illustrated in Fig. 15. The blue squares represent the ϵ values calculated among the original patterns and the P&S patterns. The orange dots represent the ϵ_c values calculated among the original patterns and the copy patterns. The mean value of ϵ calculated for the P&S patterns (blue line) are twice as larger as the mean value of ϵ_c calculated for the copy patterns (red line). Therefore, we can conclude that the suggested textured

patterns are sensitive to the copy process and can be used to distinguish the original printed document from its copy.

V. CONCLUSION

In this paper a new rich code called two level QR (2LQR) code is proposed. This 2LQR code has two levels: a public level and a private level. The public level can be read by any QR code reading application, while the private level needs a specific application with specific input information. This 2LQR code can be used for private message sharing or for authentication scenarios.

The private level is created by replacing black modules with specific textured patterns. These textured patterns are considered as black modules by standard QR code reader. Thus the private level is invisible to standard QR code readers. In addition, the private level does not affect in anyway the reading process of the public level.

The proposed 2LQR code increases the storage capacity of the classical QR code due to its supplementary reading level. Experiment results show that the storage capacity is improved by up to 28% (transition from message size equal to 272 bits to a message length of 380 bits). The storage capacity of the 2LQR code can be improved by increasing the number of textured patterns used or by decreasing the textured pattern size. All experiments show that even with a pattern size of 6×6 pixels and with an alphabet dimension $q = 8$, it is possible to obtain good pattern recognition results, and therefore a successful private message extraction. However, we are facing a trade-off between the pattern size, the alphabet dimensions and the quantity of stored information during the 2LQR code generation.

One important feature of the textured patterns used is their sensitivity to the P&S process. To take advantage of this sensitivity, we use a pattern recognition method based on maximization of correlation values among the P&S degraded versions and characterization patterns. We have tried three different types of characterization patterns: mean patterns, median patterns (for the private message sharing scenario) and original patterns (for the document authentication scenario). The mean and median characterization patterns give almost the same results of pattern detection. Therefore, either of them can be used in the private message sharing scenario. The best pattern recognition results were obtained, when the original patterns are used as characterization patterns. The original patterns can be also used for the private message sharing scenario, but in this case the blind method for pattern detection cannot be performed.

The suggested textured patterns can be distinguished only after one P&S process. Therefore, we can use the detection method with original patterns in order to ensure good document authentication results.

In our future work, we will address five different paths. The first path will concern the improvements of the pattern recognition method. The second will cover the textured pattern analysis to automate its combination process. The third will deal with message recovering and authentication attacks, such as cropping and code reconstruction. The forth path will concern the study of the second level recovery problems in the 2LQR code images captured by a camera. In the last path, the storage capacity of 2LQR code will be increased by replacing also the white modules with textured patterns, which have small density than black pixels.

REFERENCES

[1] *Information Technology—Automatic Identification and Data Capture Techniques—EAN/UPC Bar Code Symbology Specification*, ISO/IEC Standard 15420:2009, 2009.

[2] *Information Technology—Automatic Identification and Data Capture Techniques—Data Matrix Bar Code Symbology Specification*, ISO/IEC Standard 16022:2006, 2006.

[3] *Information Technology—Automatic Identification and Data Capture Techniques—Bar Code Symbology—QR Code*, ISO/IEC Standard 18004:2000, 2000.

[4] Z. Baharav and R. Kakarala, "Visually significant QR codes: Image blending and statistical analysis," in *Proc. IEEE Int. Conf. Multimedia Expo (ICME)*, Jul. 2013, pp. 1–6.

[5] C. Baras and F. Cayre, "2D bar-codes for authentication: A security approach," in *Proc. 20th Eur. Signal Process. Conf. (EUSIPCO)*, Aug. 2012, pp. 1760–1766.

[6] T. V. Bui, N. K. Vu, T. T. P. Nguyen, I. Echizen, and T. D. Nguyen, "Robust message hiding for QR code," in *Proc. IEEE 10th Int. Conf. Intell. Inf. Hiding Multimedia Signal Process. (IIH-MSP)*, Aug. 2014, pp. 520–523.

[7] 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 J. Inf. Secur.*, vol. 2014, no. 1, p. 9, 2014.

[8] T. Langlotz and O. Bimber, "Unsynchronized 4D barcodes," in *Proc. 3rd Int. Symp., ISVC 2007*, Lake Tahoe, NV, USA, Nov. 26–28, 2007, pp. 363–374.

[9] C.-Y. Lin and S.-F. Chang, "Distortion modeling and invariant extraction for digital image print-and-scan process," in *Proc. Int. Symp. Multimedia Inf. Process.*, 1999, pp. 1–10.

[10] P.-Y. Lin, Y.-H. Chen, E. J.-L. Lu, and P.-J. Chen, "Secret hiding mechanism using QR barcode," in *Proc. IEEE Int. Conf. Signal-Image Technol. Internet-Based Syst. (SITIS)*, Dec. 2013, pp. 22–25.

[11] J. Picard, "Digital authentication with copy-detection patterns," *Proc. SPIE*, vol. 5310, pp. 176–183, Jun. 2004.

[12] M. Querini, A. Grillo, A. Lentini, and G. F. Italiano, "2D color barcodes for mobile phones," *Int. J. Comput. Sci. Appl.*, vol. 8, no. 1, pp. 136–155, 2011.

[13] M. Querini and G. F. Italiano, "Facial biometrics for 2D barcodes," in *Proc. IEEE Fed. Conf. Comput. Sci. Inf. Syst. (FedCSIS)*, Sep. 2012, pp. 755–762.

[14] J. Rouillard, "Contextual QR codes," in *Proc. IEEE 3rd Int. Multi-Conf. Comput. Global Inf. Technol. (ICCGI)*, Jul./Aug. 2008, pp. 50–55.

[15] B. Sklar, *Digital Communications*, vol. 2. Englewood Cliffs, NJ, USA: Prentice-Hall, 2001.

[16] K. Solanki, U. Madhow, B. S. Manjunath, S. Chandrasekaran, and I. El-Khalil, "'Print and scan' resilient data hiding in images," *IEEE Trans. Inf. Forensics Security*, vol. 1, no. 4, pp. 464–478, Dec. 2006.

[17] M. Sun, J. Si, and S. Zhang, "Research on embedding and extracting methods for digital watermarks applied to QR code images," *New Zealand J. Agricult. Res.*, vol. 50, no. 5, pp. 861–867, 2007.

[18] I. Tkachenko, W. Puech, O. Strauss, J.-M. Gaudin, C. Destruel, and C. Guichard, "Fighting against forged documents by using textured image," in *Proc. 22th Eur. Signal Process. Conf. (EUSIPCO)*, Sep. 2014, pp. 790–794.

[19] R. Ulichney, *Digital Halftoning*. Cambridge, MA, USA: MIT Press, 1987.

[20] 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 *Proc. SPIE*, vol. 6505, p. 65051T, Feb. 2007.

[21] R. Villán, S. Voloshynovskiy, O. Koval, and T. Pun, "Multilevel 2-D bar codes: Toward high-capacity storage modules for multimedia security and management," *IEEE Trans. Inf. Forensics Security*, vol. 1, no. 4, pp. 405–420, Dec. 2006.

[22] S. V. Voloshynovskiy, O. Koval, F. Deguillaume, and T. Pun, "Visual communications with side information via distributed printing channels: Extended multimedia and security perspectives," *Proc. SPIE*, vol. 5306, pp. 428–445, Jun. 2004.

[23] S. Vongpradhip and S. Rungraungsilp, "QR code using invisible watermarking in frequency domain," in *Proc. IEEE 9th Int. Conf. ICT Knowl. Eng. (ICT & Knowledge Engineering)*, Jan. 2012, pp. 47–52.

[24] L. Yu, X. Niu, and S. Sun, "Print-and-scan model and the watermarking countermeasure," *Image Vis. Comput.*, vol. 23, no. 9, pp. 807–814, 2005.

Iuliia Tkachenko received the M.S. degree in applied mathematics from Dnipropetrovsk National University, Ukraine, in 2010, and the M.S. degree in cryptography and information security from University Bordeaux 1, France, in 2012. She is currently pursuing the Ph.D. degree with the Laboratory of Informatics, Robotics and Microelectronics, France, with the collaboration of Authentication Industries. Her research interests include multimedia security, document authentication, and cryptography.
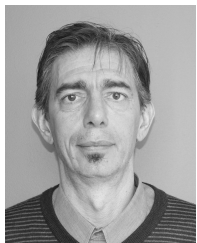
William Puech received the Diploma of Electrical Engineering from the University of Montpellier, France, in 1991, and the Ph.D. degree in signal-image-speech from the Polytechnic National Institute of Grenoble, France, in 1997. He started his research activities in image processing and computer vision. He served as a Visiting Research Associate with the University of Thessaloniki, Greece. From 1997 to 2000, he was an Assistant Professor with the University of Toulon, France, with research interests including methods of active contours applied to medical images sequences. From 2000 to 2008, he was an Associate Professor and has been a Full Professor of Image Processing with the University of Montpellier, since 2009. He is with the Laboratory of Computer Science, Robotic and Microelectronic of Montpellier. He is the Head of the ICAR team (Image & Interaction). He has published over 15 journal papers, eight book chapters, and more than 80 conference papers. His current interests are in the areas of protection of visual data (image, video, and 3-D object) for safe transfer by combining watermarking, data hiding, compression, and cryptography. He has applications in medical images, cultural heritage, and video surveillance. He is an Associate Editor of the *Journal of Advances in Signal Processing* (Springer), *Signal Processing: Image Communications* (Elsevier), and *Signal Processing* (Elsevier), and a Reviewer for more than 15 journals (the IEEE TRANSACTIONS ON IMAGE PROCESSING, the IEEE TRANSACTIONS ON MULTIMEDIA, the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, the IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY, *Signal Processing: Image Communication*, *Multimedia Tools and Applications, etc.*) and for more than ten conferences (the IEEE ICIP, EUSIPCO, etc.).

Christophe Destruel received the Engineering degree from the University of Toulouse, France, in 1995. He worked for ten years in collaboration with the French Spatial Agency on research programs related to its main fields of interest, computer graphics and image processing, to develop the use of 3-D paradigm in space imaging. He is the Scientific Director of the start-up Authentication Industries that proposes innovative solutions to authenticate valuable documents. Respecting industrial constraints, his contributions have been published in various international conferences.

Olivier Strauss received the Ph.D. degree in signal processing from Montpellier University, France, in 1991. He is an Associate Professor of Signal and Image Processing with the University of Montpellier. He obtained the accreditation to supervise research in 2008. His current research interests include signal and image processing, computer vision, medical imaging, uncertainty management, advanced statistics, possibility, and imprecise probability theories. He is currently the Head of the Robotics Department and a member of the Scientific Council with Montpellier University.

Jean-Marc Gaudin is one of the founders of Authentication Industries and one of the inventors of the AiCode.

Christian Guichard received the M.B.A. degree in 1987. He is the Chairman of the company Authentication Industries, and is well known as an expert in the fight against fraud and counterfeiting throughout 95 countries, after nearly 20 years of experience creating an international network of partners and specialized printers. Since the inception of the company in 2012, he personally invested over €550 000. The other founders have nearly 20 years in the field of security, involving other technologies, such as biometrics, RFID, nanotechnology, and cryptography. These complementary and experienced founders, with the help of C. Destruel, AI's Scientific Director, launched an ambitious R&D program to accelerate the development of their innovative patented solution: the AiCode.