International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 07 Issue: 07 | July 2020 www.irjet.net p-ISSN: 2395-0072

A Secure QR Code from Multi-Layer Encryption System

Sabitha Sathyan[1], Husna Mol[2], Pratheksha P.R.[3], Deepthy S[4]

[1,2,3]B.Tech Student, Computer Science and Engineering, APJ Abdul Kalam Technological University, Kerala, India
[4]Asst. Professor, Computer Science and Engineering, Mount Zion College Of Engineering, Kadammanitta, Kerala, India

---

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 (2LQR), 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.*

***Key Words*: QR Code, Two storage levels, Private message, Document authentication, pattern recognition, print –and- scan process.**

1. INTRODUCTION

Today, graphical codes, such as EAN-13 barcode Quick Response (QR) code Data Matrix 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 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 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 pattern scan 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 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.

2. RELATED WORKS

M. Querini, A. Grillo, A. Lentini, and G. F. Italiano. 2D color barcodes for mobile phones. IJCSA, 8(1):136– 155, 2011.

They propose a new high capacity color barcode, named HCC2D (High Capacity Colored 2Dimensional), which use colors to increase the barcode data density. The introduction and recognition of colored modules poses some new and non-trivial computer vision challenges, such as handling the color distortions introduced by the hardware equipment that realizes the Print &Scan process. We developed a prototype for generating and reading the HCC2D code format, both on desktops (Linux and Windows platforms) and on mobile phones (Android platforms). We tested this prototype in many experiments considering different operating scenarios and data densities, and compared it to known 2dimensional barcodes.

B. Sklar. Digital communications, volume 2. Prentice Hall NJ, 2001.

Exceptionally accessible, this book presents the often "difficult" concepts of digital communications in an easy-to-understand manner— without diluting the mathematical precision. Using a student-friendly approach, it develops the important techniques in the context of a unified structure (in block diagram form)—providing organization and structure to a field that has, and continues, to grow rapidly, and ensuring that students gain an awareness of the "big picture" even while delving into the details (the most up-to-date modulation, coding, and signal processing techniques that have become the basic tools of our modern era). It traces signals and key processing steps from the information source through the transmitter, channel, receiver, and ultimately to the information sink**.**

3. PROBLEM DEFINITION

This section is split into three sub-sections (3.1, 3.2, 3.3). We start with a description of the standard QR code features. Finally, to highlight authentication of physical documents, the distortions added to any image during the P&S process.

3.1 QR Code Features

The QR code was invented for the Japanese automotive industry by Denso Wave1corporation 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. 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 [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%.

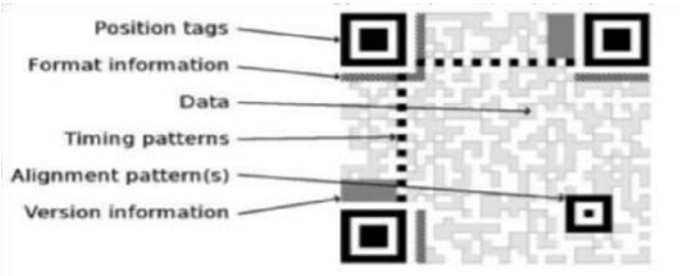

Fig 3.1: *Specific QR Code Structure***.**

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. 3.1, 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 code words, application of mask pattern, placement of code words and function patterns into the QR code. The QR code recognition algorithm includes the scanning process, image binarization, geometrical correction and decoding algorithm.

3.2 Rich Graphical Codes

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 applications2. However, most of these generators prefer to sacrifice the possibility of error correction for attractive design. 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 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 , 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.

3.3 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.

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. The nonuniform distribution of toner is the manufacturer defect and introduces geometrical image displacement. 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. 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. 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, where is the parameter of a power function, which represents an intensity to voltage response curve. This process introduces non-linear transformation. All these changes mentioned produce pixel value distortions. 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.

4. 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.

4.1 Two level QR (2LQR) Code Generation

Like the standard QR code, the 2LQR code has the same specific structure, which consists of position tags, alignment patterns, timing patterns, version and format patterns. 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 qary (q 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 Storage in the proposed 2LQR code. We suggest to use the 2LQR code for two scenarios: for private message sharing and for document authentication. In private message sharing scenario, the black modules of these patterns are also replaced by textured patterns.

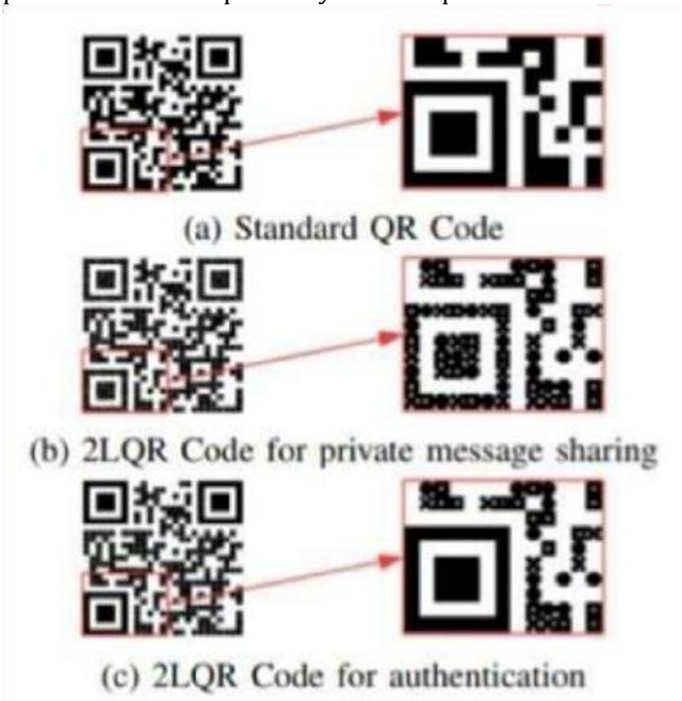

Fig 4.1: *Comparison*

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 and then are placed in the position tags. These textured patterns will be used for pattern detection. The reading algorithm can decode the private message M_{priv}, if the permutation and the ECC algorithm used for message encoding are known. 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 code words. Then, the choice of error correction level is performed and the error correction code words using the Reed-Solomon code are generated. After that, the data and error correction code words are arranged in the correct order. In order to be sure that the generated QR code can be read correctly.

4.1.1 Textured Pattern Selection

The textured patterns P_i, i = 1,...,q are images of size p×p pixels. We choose q patterns from a database of Q >> 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 Si,i = 1,...,q be the P&S degraded versions of textured patterns Pi,i = 1,...,q. The Pearson correlation between a pattern Pi and a P&S pattern. 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 could be used in 2LQR code generation.

1) Each textured pattern Pi,i = 1,...,q has to be better correlated with its P&S degraded version Si,i = 1,...,q than with all other P&S degraded versions Sj,j = 1,...,q,i 6= j: i,j {1,...,q},i 6= j,cor(Pi,Si) > cor(Pi,Sj).(2)
2) The P&S degraded version Si,i = 1...q of each pattern has to be better correlated with its original pattern Pi,i = 1...q than with all other original patterns Sj,j = 1...q,i 6= j: i,j {1,...,q},i 6= j,cor(Pi,Si) > cor(Pj,Si) .
3) These two conditions (2) (3) can be rewritten in the form: i {1,...,q},cor(Pi,Si) = max j {1,...,q} (cor(Pi,Sj)) = max j {1,...,q} (cor(Pj,Si)).
4) The condition (4) is valid for all values of j (j {1,...,q}), but if the value j = i is excluded, the condition (4) can be rewritten in the form: i,j {1,...,q},cor(Pi,Si) = max i6=j (cor(Pi,Sj)) + i1 = max i6=j (cor(Pj,Si)) + i2.
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 threshold. That is why the limits for distances i1 and i2 are set: i1 and i2 ,
6) the theoretical distance values have to be in the interval i1,i2 [,2]. Therefore, only the textured patterns, which respect the conditions (4) (6), can be combined, and used for 2LQR code generation.

The code word C_{priv} is inserted in standard QR code by replacing the black modules with textured patterns P1,...,Pq respecting the code word 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 , see Fig. 2.b. In the case of authentication scenario, the standard position tags keep unchanged black modules.

4.2 Storage capacity of 2LQR code

In this section the storage capacities of proposed 2LQR code. Let N2 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 N2/2 is approximately the number of black modules in standard QR code.

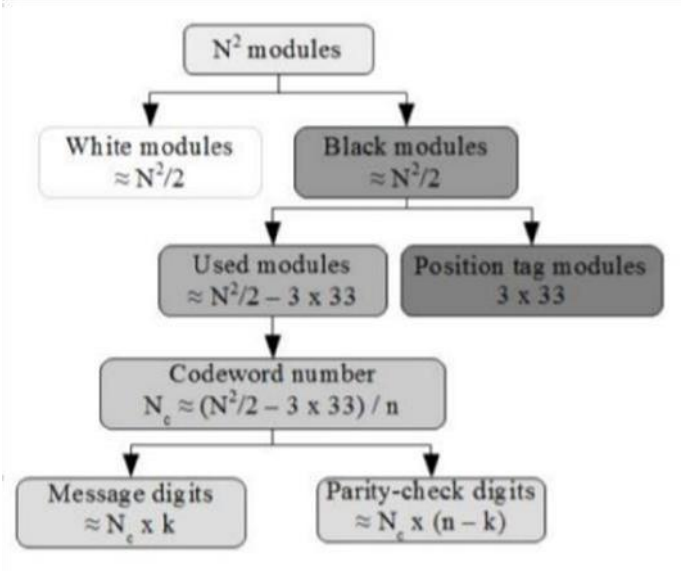

Fig 4.2: *Storage Capacity OF 2LQR Code of Size*

Let n be the total number of digits in a code word, k be the number of message digits, n k be the number of error correction bits in a code word. Therefore, the number of code words, that could be inserted in the second level of a 2LQR code, is approximately equal to Nc (N2/2 3 × 33)/n. And the number of message digits is approximately equal to Nc ×k, that is why the length of message is approximately equal to log2(q)×Nc ×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 N2 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.

4.3 Recognition Method

The QR code reproduction implies the use of the printing process and the scanning process. First, the geometrical distortion of P&S 2LQR code has to be corrected during the preprocessing 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 ×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×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. We compare the P&S patterns with the original patterns, that is why the characterization patterns are replaced by: The last steps of the 2LQR code reading process are unscrambling using key K and ECC decoding of the obtained code word $C0_{priv}$. 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}.

5. EXPERIMENTAL RESULTS

This section illustrates both the generation steps of the 2LQR code and the message extraction steps. Then, the storage capacities of the 2LQR code is discussed. 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 Go lay code [11,6,5]. Each code word 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 Go lay code [11,6,5] we can store 114 ternary digits, that corresponds to nearly 180 message bit.

5.1 2LQR code generation

The 2LQR code generation consists of four steps: standard QR code generation, code word generation, pattern selection and replacement of black modules in QR code. Standard QR code generation. The standard QR code with public message Mpub "John Doe - 13/05/1958 - New York" is generated by using a free online QR code generator. The generated standard QR code, version V 2, is illustrated. The actual size of this QR code is 1.2×1.2 cm².

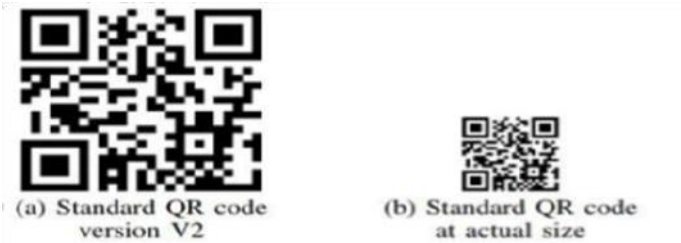

Fig 5.1: *The example of a) Standard QR code with public message M_{pub}, b) Standard QR code at actual size defined at 600 dpi, 1.2×1.2 cm²*

The original and the P&S degraded versions of these patterns satisfy conditions (4) (6) with = 0.25.

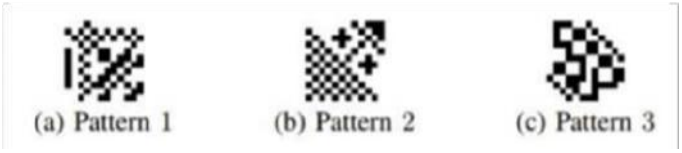

Fig 5.1.1: *The three textured patterns used for private level generation: a) Pattern 1, b) Pattern 2, c) Pattern 3.*

5.2 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).

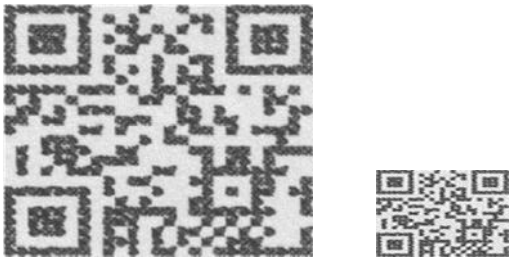

Fig 5.2: *The example of a) P&S 2LQR code for private message sharing and b) P&S 2LQR code at actual size defined at 600 dpi.*

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.

6. METHODOLOGY

Coding Language : JAVA
Database : MY SQL

7. 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 3.3, 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. 7.1. Due to the P&S impact, it is difficult to model the P&S degraded versions of proposed textured patterns.

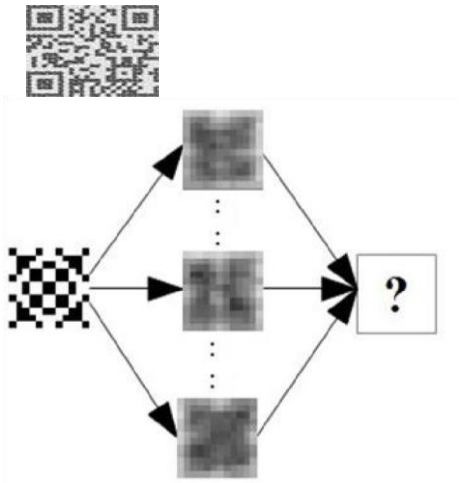

Fig 7.1: *Examples of modifications added to textured patterns during the P&S process.*

We decided to measure the difference between the document after a P&S process and a copy of the document after 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 7.1.1.

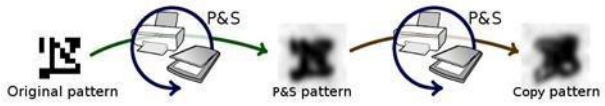

Fig 7.7.1: *Production of P&S pattern and copy pattern from the original pattern.*

8. CONCLUSION

In this paper a new rich code called two level QR (2 LQR) 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 fourth 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] D. Kornack and P. Rakic, "Cell Proliferation without Neurogenesis in Adult Primate Neocortex," Science, vol. 294, Dec. 2001, pp. 2127-2130, doi:10.1126/science.1065467.

[2] M. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.

[3] R. Nicole, "Title of paper with only first word capitalized," J. Name Stand. Abbrev., in press.

[4] K. Elissa, "Title of paper if known," unpublished.

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

[6] T. V. Bui, N. K. Vu, T. T.P. Nguyen, I. Echizen, and T. D. Nguyen. Robust message hiding for QR code.

[7] A. T. P. Ho, B. A. M. Hoang, W. Sawaya, and P. Bas. Document authentication using graphical codes.