Development of High Data Capacity Color Code

K. S. Sim, Z. Y. Lim, and F. F. Ting

***Abstract*—This paper presents the implementation of high data capacity code by embedding multiple colors into the code. The main intention of this project is to implement a code that consists at least double the data capacity as compared to the present Quick Response (QR) code. This paper explains the elements that consist in the proposed high data capacity code layout. The high data capacity code utilizes color multiplexing technique to represent data by using 8 kinds of colors. The colors involved are three primary colors, three secondary colors, black and white. Besides, the code is embedded with same error correction algorithm as QR code namely Reed Solomon Error Correction. In this project, a decoder application is developed on personal computer to decode the information from the camera-captured code. The results show that the developed decoder software is capable to perform decoding without any error within the captured range of 7 cm to 15 cm. Since the developed second prototype consist of similar number of module with QR Version 8, the performance is assessed to compare with QR Version 8. As outcome of this project, the developed high data capacity code is to achieve more than doubles the data capacity of QR code Version 8.**

***Index Terms*—High, data, capacity, color, code.**

I. Introduction

In 1948, Bernard Silver and Norman Joseph Woodland had invented barcode to store product information [1]. The barcode is a single dimensional (1D) optical data representation in parallel lines format. The data are stored systematically by manipulating the size and width between the parallel lines. An optical scanner is specially designed to decode the barcode. As the advancement of telecommunication and ICT, smartphones have become common in the market. Thus, some applications have been developed to allow mobile phone to read the data of the barcode easily by capturing the barcode image using the smartphone's camera.

However, as technology advances, the size required to store information keep expanding. Hence, barcode has reached the bottleneck where its maximum size can only store up to 48 alphanumeric character. The single dimensional (1D) barcode eventually has improvised into two dimensional (2D) barcode. There are several types of 2D barcode that have been implemented. However, the 2D barcode that is most commonly utilized nowadays is the Quick Response (QR) code.

The QR code is a trademark for a type of matrix barcode. It was developed in 1992 by Denso Wave Inc. to serve the purpose of tracking the vehicles during manufacturing with a requirement of high decoding speed [2]. The colors involved in the QR code are black and white colors [3]. The data that represented in black or white depends on the value of bit [3]. The design of QR code is comprised by two major elements. The first element is known as function patterns. The function patterns involve finder pattern, timing pattern and alignment pattern. These patterns that function as the reference points is employed to pinpoint the exact position of encoded information on the QR code. The second element is known as encoding region. The encoding region stores the information such as data format, version of the code, error parity data and actual data. According to the format and version of the QR code, the decoder will alter the decoding algorithm accordingly in order to decode the relevant information from the QR code [4].

There are few major innovations that make the QR code becomes a breakthrough in this barcode field. First, QR code is able to achieve more than 100 times data capacity as compared to 1D barcode. Besides, instead of storing solely alphanumerical, QR code is able to store numerical, alphanumerical, binary data and Japanese words (Kanji) as well. In this way, the data capacity of the code can be maximized efficiently. Next, with the contribution of the error correction and function patterns of the QR code, the QR code is readable in any direction of alignment. Last but not least, the greatest attribute of QR code is the error correction algorithm embedded in the QR code. The error correction algorithm employed by QR code is known as Reed-Solomon error correction. The error correction level can be customized range from 7% to 30%.

Nowadays, QR code becomes extremely popular and can be found everywhere. At first, the purpose of creating QR code is to provide more data capacity to store product information for inventory tracking. However, after the smartphones emerged, the QR code is widely used for all kind of purposes due to its convenience and efficiency. In present day, the QR code that can be found is utilized in ticketing, entertainment, commercial tracking, and even product marketing. Since QR code can encrypt all kinds of information, many actions can be performed by just scanning the QR code using the smartphone camera. In less that one second, the smartphone can automatically add contacts, browse website, collect information and even do transaction. Besides, QR code is also used in education to increase the efficiency in teaching [5], [6]. Hence, QR code improve our lifestyle by increasing the efficiency of various task in our daily life.

In present, the QR code has evolved to Version 40. The maximum data storage capability of QR code Version 40 is approximately 4 kilobytes including the error correction information, which means that the actual data capacity is around 3 kilobytes [7]. In this technology era, the technology keeps on updating. Hence, the demand on the QR code is emerging too. Some may demand that the QR code is able to store a picture, a word document or a short video within the

Manuscript received April 17, 2018; revised July 5, 2018. This work was supported by the Multimedia University under Grant MMUI/160151.

The authors are with the Faculty of Engineering and Technology, Multimedia University, Malaysia (e-mail: kssim@mmu.edu.my, limzhengyou@gmail.com, sicily.ting@gmail.com).

doi: 10.18178/ijmlc.2018.8.4.709

code. In this way, they are able to access the image, document or video with a single scanning without accessing the internet. However, the current QR code does not have sufficient data storage to perform these tasks.

In this project, the main objective is to develop a code that is able to store twice as much data than QR code with the same module size. In order to accomplish this objective, this project proposes the high data capacity code which encrypts the information with 8 colors in square block. The layout of the code is in rectangular form. Since the QR code consists of the high efficiency data error correction feature, the proposed high data capacity code is also designed by considering Reed Solomon Error Correction as error correction algorithm. For the second prototype, the color code is designed according to the number of module consisted in QR Code Version 8. In result, the performance of the developed high data capacity code is assessed by comparing with the data capacity of QR Code Version 8. Beside developing the high data capacity code, another objective of the project is to develop the decoder application to decode the high data capacity code efficiently.

II. Design and Implementation

A. Layout Design

Fig. 1 shows the designed layout of the High Data Capacity Code.The feature of each elements of the High Data Capacity Code is elaborated as shown in Table I.

TABLE I: Feature of Elements in High Data Capacity Code

Component Name	Function
Quiet Zone	Margin in white color between the reference boundary and data encoding region. This zone distinguish the reference boundary and data encoding region. Used to extract out the data encoding region during decoding process.
Reference Boundary	Boundary of the code which makes up of cyan, yellow, magenta and black colors. The Reference Boundary is utilized to track the location of the code. The colors involved in the Reference Boundary provide reference value for image processing during decoding.
Reference Alignment Pattern	Three blocks on the corner of the code: upper left, upper right and lower right. Reference Alignment Pattern is the element that utilized to determine whether the code is misaligned. The colors involved in the Reference Alignment Pattern provide reference value for image processing during decoding.
Data Encoding Region	Region where the actual information and error correction codewords are stored into color block format. The data is encoded as Reed Solomon codeword blocks.

B. Reed Solomon Correction Code

QR Code error correction feature is executed by embedding the Reed-Solomon Correction Code to the actual information. Reed-Solomon Code is a mathematical error correction method that is prominently employed in consumer technologies such as CD, DVD, Blu Ray Discs etc [8].

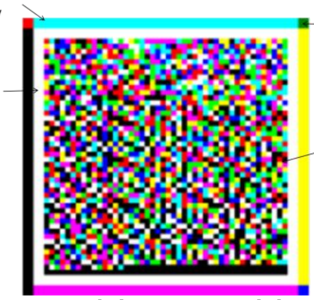

Fig. 1. High data capacity code layout.

This algorithm was initially implmented to reduce the noise in the communication electronic field. This technique is able to perform correction at the byte level, and is highly relevant to deal with concentrated burst errors [9].

The Reed Solomon codeword block is generated by utilizing polynomial mathematical formula as shown in (1).

$$c(x) = g(x) \cdot i(x) \quad (1)$$

where $g(x)$ is the generator polynomial and $i(x)$ is the message block.

Thus, the number of bits per codeword block is determined by the number of bits involved in one symbol. The calculation is shown in (2).

$$\text{BlockSize} = n * (2^n - 1) \quad (2)$$

where BlockSize is the total number of bits per codeword block and n is the number of bits involved in one symbol. For example, the maximum value of the all the symbols is 4, the number of bits (n) is 2. Thus, the block size will be 6 bits.

Then, the message size which is the actual data to be encoded depends on how many parity bits reserved as Reed Solomon correction bits. The calculation is as shown in (3).

$$\text{MessageSize} = \text{BlockSize} - \text{ParitySize} \quad (3)$$

where $MessageSize$ is the number of bits representing the actual information and $ParitySize$ is the number of bits representing the Reed Solomon parity bits for error correction purpose.

Since ASCII table consists of 256 data, thus the high data capacity code is designed with the capability to encode 256 number of data. Hence, 256 data requires 8-bit binary number to represent a single character (n=8). The size of Reed Solomon is determined by the number of bits (n) as shown in (2). Thus, the block size is equal to 255 bits. Then, the parity size is set to 75 bits, which is approximately 30% of the block size. Thus, the actual message size is calculated by using (3), which is 180.

C. Color Multiplexing

During encoding process, the data are encoded into three different layers known as top layer, middle layer and bottom layer that consists of 2116 square blocks per layer. In each layer, the working principle is similar to QR code, where the value of bit is represented by black and white square blocks. Bit with value ‘1’ is assigned with white color. Vice versa, bit with value ‘0’ is assigned with black color.

Then, all of the layers of square blocks are transformed into color square blocks by implementing the color multiplexing algorithm.

The color multiplexing algorithm is applied by utilizing the additive color model approach [10], [11]. The additive color model involved in this technique is the three primary color known as Red, Green and Blue.

Thus, every layers are replaced by different colors. Red color denotes the top layer, green color denotes the middle layer and blue color denotes the bottom layer.

By mixing two of the primary colors on the specific square block that represented by different layers, secondary colors are produced namely Yellow, Magenta and Cyan colors. However, when three primary colors present on the same specific square block in all the layers, white color will be assigned. Vice versa, if all three layers do not consists any color on the specific square block, black color will be assigned [13]. The layer multiplexing technique is shown in Table II.

TABLE II: LAYER COLOR MULTIPLEXING

Top Layer	Middle Layer	Bottom Layer	Resulting Color
0	0	0	White
1	1	0	Yellow
1	0	1	Magenta
1	0	0	Red
0	1	1	Cyan
0	1	0	Green
0	0	1	Blue
0	0	0	Black

By using the multiplexing technique, it is capable to develop a high data capacity code with multiple colors that involves a diversity of 8 colors, which are white, black, cyan, magenta, yellow, red, green and blue. Fig. 2 (a), (b) and (c) illustrate the information represented in monochrome in each layer: top layer, middle layer and bottom layer. Fig. 2(d) illustrates the resulting colored code.

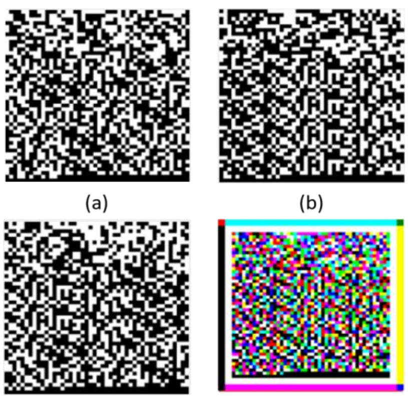

Fig. 1. (a) Top Layer (b) Middle layer (c) Bottom layer (d) Resulting color code.

D. Decoder Application

Lastly, a decoder application is developed in order to perform image processing and read the information from the captured high data capacity code.

In the beginning, the decoder algorithm begins with detecting the reference alignment pattern of the multi color code. Then, according to the detected alignment, auto realignment is performed to assure the multi color code is on upright position. After that, the color value of each block of the reference alignment pattern is determined and stored as reference for further image processing.

Then, the algorithm is continued by detecting the boundary of the high data capacity code and obtain the data encoding region out from the quiet zone. During this process, the colors' value are also determined and stored as reference. Next, the color blocks that represent the data is enhanced with some image processing techniques. These developed techniques are used to enhance the image. In this way, it is able to decrease the error rate during the decoding process.

Next, the algorithm proceed to slice the colored layer back into top, middle and bottom layers that only involves monochrome colors. Meanwhile, the error correction Reed Solomon algorithm is carried out to detect error and perform the error correction. Finally, all three layers are decoded into binary bits data. Then the alphanumerical data is constructed according to the binary bits data. The flowchart of the algorithm is as illustrated in Fig. 3.

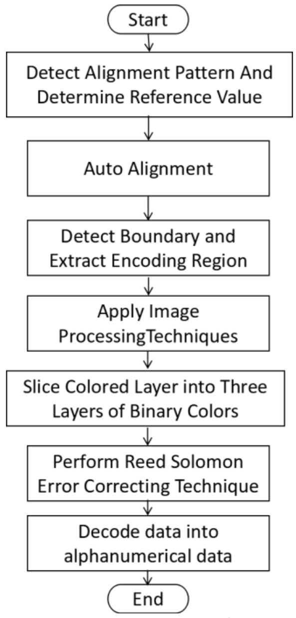

Fig. 2. Decoder application flowchart.

III. DESIGN AND IMPLEMENTATION

A. Experimental Setup

In this project, two software applications are implemented which is the encoder application and decoder application. The Integrated Development Environment (IDE) that utilized to develop the applications is Microsoft Visual Studio. This IDE is programmed by using C# language [12]. The application is executed on a workstation with Windows operating system.

The performance of the developed decoder application is assessed by experimental procedure. At first, the high data capacity code is adjusted to the width and height of 15 mm × 15 mm. Then, the high data capacity code is printed on a white 80g/m2 density A4 paper.

Next, a mobile phone Asus Zenfone 4 is used to capture the printed high data capacity code for each specified distance. The resolution of the rear camera is 16 megapixels. The high data capacity code images are captured under the condition of indirect daylight illuminance. 5 images are captured for each specified distance using the rear camera of the mobile phone. Thus, experimental sets is comprised of 85 images.

B. Decoder Performance

The performance of the decoder application is assessed by calculating the average decode error rate. Thus, the average decode error rate is defined as number of unsuccessful decoding rate among the five captured images for each pre-defined range between the camera and the image. Fig. 4 illustrates the performance of the decoder in term of average error rate on each pre-defined distance.

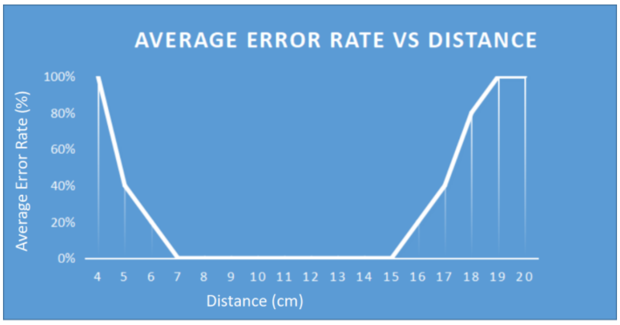

Fig. 3. Average decoding error rate vs captured distance.

C. Data Capacity of High Data Capacity Code

The performance of the high data capacity code is assessed by benchmarking with the data capacity of QR code. The information encoded in QR Code is in module form. The module is defined as the smallest black or white square block on the QR code.

In this project, the second prototype of the high data capacity code consists similar number of module with QR Code Version 8. Thus, the data capacity of the high data capacity code is associated with the Version 8 of QR code. The comparison result is shown in Table III.

TABLE III: COMPARISON BETWEEN QR CODE (VERSION 8) WITH HIGH DATA CAPACITY CODE (SECOND PROTOTYPE)

Code	Module (Square Block)	Binary Data Capacity
QR Code (Version 8)	49*49 = 2401	192
High Data Capacity Code	46*46 = 2116	540

By referring to the results shown in Table III, the high data capacity code is capable to encode 540 binary data, while Version 8 of QR code is able to store 192 binary data. Thus, the developed high data capacity code is having 2.81 times more data capacity than QR code Version 8. Thus, it shows that the high data capacity code is able to store more than doubles the data capacity of QR code.

IV. CONCLUSION

According to the result, it indicates that the developed decoder application is able to decode the implemented high data capacity code without any error within the captured distance from 7 cm to 15 cm. It is found that the quality of the captured image is the main factor. Hence, this project also implements some image enhancement and processing techniques in order to uplift the performance of the decoder. These image processing techniques have been patented (PI 2017705178) [14].

REFERENCES

[1] D. Davis and P. Soumya, "Enhanced security on 2D barcode using picture embedding picode approach," *Imperial Journal of Interdisciplinary Research*, vol. 3, no.6 , pp. 112-115, 2017.

[2] J. A. Lin and C. S. Fuh, "2D barcode image decoding," *Hindawi*, vol. 1, 2013.

[3] QR codes: 'Information technology. Automatic identification and data capture techniques', QR Code 2005 Bar Code Symbology Specification, January 2007.

[4] I. C. Dita, M. Otesteanu, and Q. Franz, "Data matrix code — A reliable optical identification of microelectronic components," in *Proc. the IEEE 17th Int. Sym. for Design and Tech. in Electronic Packaging*, October 2011.

[5] R. Dorado, E. Torress, and C. Rus, "Mobile learning: Using QR codes to develop teaching material," *IEEE-Technologies Applied to Electronics Teaching*, June 22-24, 2016.

[6] K. Y. Lin, Y. C. Yang, and C. Y. Lai, "The use of quick response codes for teaching pharmacology to college nursing students in Taiwan," *Computers Informatics Nursing*, vol. 35, no. 3, pp. 152-157, Mar. 2017.

[7] H. Kato and K. T. Tan, "Pervasive 2D barcodes for camera phone applications," *IEEE Pervasive Computing*, vol. 6, pp. 76–85, Oct.-Dec. 2007.

[8] D. Omerasevic, N. Behlilovic, and S. Mrdovic, "An implementation of secure key exchange by using QR codes," in *Proc. 56th International Symposium ELMAR-2014*, September 2014.

[9] B. S. Eli, B. T. Iddo, G. Ariel, and R. Micheal, "Improved concrete efficiency and security analysis of Reed-Solomon PCPPs," *Electronic Colloqium on Computational Complexity*, vol. 1, no. 73, 2016.

[10] S. Vongpradhip, "Use multiplexing to increase information in QR-Code," in *Proc. 8th International Conference on Computer Science Education*, pp. 361-364, April 2013.

[11] P. S. Andre and R. A. S Ferreira, "Colour multiplexing of quick-response (QR) codes," *Electronics Letter*, vol. 50, no. 24, pp. 1828-1830, Nov. 2014.

[12] Microsoft Visual C#. Microsoft, 1989. [Online]. Available: http://msdn.microsoft.com

[13] Z. Y. Lim, K. S. Sim, and E. K. Wong, "High capacity multi colored code system," in *Proc. International Conference on Robotics, Automation and Sciences*, November 2017.

[14] K. S. Sim and Z. Y. Lim, *Image Reconstruction of Camera Captured Microscale Multi-Colored Object*, 2017.

Sim Kok Swee is a professor in the Faculty of Engineering and Technology (FET) at the Multimedia University, Malacca, Malaysia. His main areas of research are IC failure analysis, application of SEM, noise quantization, biomedical engineering and image processing. He has been working in the industrial and teaching line for more than 25 years. Ir. Prof Dr. Sim Kok Swee is an associate fellow for Malaysia Academic Science Malaysia, senior panel for Engineering Accreditation Council and Malaysian Qualifications Agency (MQA), fellow member of The Institution of Engineering and Technology (IET), fellow member of The Institution of Engineers, Malaysia (IEM), senior member of Institute of Electrical and Electronics Engineers (IEEE). He is heavily involved in IEM and the Institution of Engineering and Technology (IET) as MMU student chapter advisor.

Lim Zheng You is a graduate research assistant in Multimedia University, Malaysia. He holds BEng. (Hons) electronics majoring in robotics and automation with First Class Honour in Multimedia University, Malaysia in 2016. Mr Lim is the Merit Winner recipient in APICTA International 2017 in Dhaka, Bangladesh. His current researches are image processing and brainwaves signal processing. He is a graduate member of Board of Engineer Malaysia.

Ting Fung Fung is a graduate research assistant in Multimedia University, Malaysia. She holds master in engineering science in 2017 and BEng. (Hons) electronics majoring in robotics and automation in Multimedia University, Malaysia in 2015. Her current research are image processing. She is a graduate member of Board of Engineer Malaysia.