Performance Analysis of Binary Watermarking Algorithm Against Non-Geometric Attacks

Zoran Veličković, Dejan Blagojević, Marko Veličković
Akademija tehničko-vaspitačkih strukovnih studija, odsek Niš
Niš, Srbija
zoran.velickovic@akademijanis.edu.rs, dejan.blagojevic@akademijanis.edu.rs, marko.velickovic@akademijanis.edu.rs

Abstract— **In this paper, the performance of the recommended algorithm for color image protection by inserting a binary watermark - QR code into the chrominance channel of the carrier image was determined. Algorithm performance is determined in cases of non-geometrics attacks such as JPG compression, Gaussian noise, Speckle noise and Salt & Pepper. For certain attack intensities, differential threshold values were determined when the watermark can be extracted without errors. Values of a certain differential threshold do not cause significant degradation of image quality, and compared to the degradation caused by individual attacks, they are significantly smaller. The information encoded by the QR code can be used to identify the author and source after decoding. The results obtained in the experimental part of the work confirm the high resistance of the proposed algorithm to non-geometrics attacks. In the continuation of the research, the performance of the algorithm for geometric attacks will be determined and the ability of the QR code to correct errors in extracted watermarks will be tested.**

Keywords- Watermarking; YCbCr color model; non-geometrics attacks; QR code; binary watermark; NCC; Bit error.

I. INTRODUCTION

In the modern world, the exchange of digital content is at an all-time high [1]. The exchange of multimedia content on social networks is especially intense. With the help of ubiquitous tools, the manipulation of digital multimedia contents is simple, they can be easily downloaded from the network, modified, and then illegally distributed for commercial, propaganda or other purposes [2]. The problem that is increasing every day is related to the protection of copyrights, that is, to the protection of multimedia contents against copying and illegal use. On the other hand, unwary users may be offered tendentious content originating from unreliable and unverified sources. Users who do not pay attention to the sources of multimedia content can easily be misled or deceived. The most recent examples of misuse of multimedia content that have led to public disturbance have been caused by unverified information related to the COVID 19 pandemic, that is, to the war conflict in Ukraine. In [3], a secure technological concept for accessing official content and sharing it on the Web is described.

In order to prevent the illegal use of multimedia content, that is, to recognize the source of the content, numerous techniques have been developed [4], [5]. Among other things, multimedia content (music, images, video, ...) can be protected by the classic PKI (Public Key Infrastructure) encryption technique, but for several reasons, they are not adequate for this type of content. Insertion of invisible (secret) information - a watermark into multimedia content (watermarking) is a protection technique that solves all the problems of classic multimedia content encryption techniques. For this work, it is significant that the proposed watermarking insertion technique does not require a complex PKI infrastructure. Although the proposed protection technique can be applied globally, it is very suitable for implementation in in-house solutions.

A color image, grayscale image or binary image can be used as a watermark. In this work, a binary image representing coded information using one of the QR codes was used as a watermark. Therefore, the watermark can be invisibly inserted into the supporting image with the proposed algorithm, while the watermark can be extracted from the protected image with the inverse algorithm, with known insertion parameters. By extracting the watermark and then decoding it, ownership of the multimedia content can be proven, that is, the real source of information can be decoded.

In general, watermark insertion algorithms should not cause noticeable degradation of the supporting images. At the same time, the extraction algorithms should provide reliable extraction of the watermark, that is, the QR code that usually encodes information about the author, owner or source of the content. These are two conflicting technical requirements that insertion/extraction algorithms need to resolve. The mathematical apparatus used in this work for inserting and extracting the watermark in the carrier image is based on YCbCr transformation [6], block SVD (Singular Value Decomposition) [7] and QR coding [8]. To reduce visible image degradation, watermark insertion is performed in the Cr chrominance channel using block SVD.

If the watermark is removed from the protected image, the copyright may be compromised or even the source of the information may be falsified. Good algorithms should be resistant to attempts to remove the watermark from the image. In practice, geometric and non-geometric attacks are used to remove a watermark from a protected image. Non-geometric attacks are based on digital image processing and consist in adding interference that degrades protected carrier images but also complicates watermark extraction. This paper analyzes the performance of the recommended algorithm in relation to standard non-geometric attacks. As for geometric attacks, they are based on cutting off part of the image, rotating or scaling the image and will be discussed in the continuation of the research. The performance of the proposed algorithm was measured by the quality of the extracted watermark for several standard

attacks. Number of bit errors and NC (Normalized Correlation Coefficient) were used as objective parameters for measuring the quality of the extracted watermark. In the experimental part of the work, the well-known color pictures Lena, Mandril, Pepper, Airpline and Fruits were used as carriers. The proposed algorithm consists of an insertion part and a watermark extraction part. Given that the proposed algorithm allows the selection of a differential threshold T, it is varied within limits that do not cause significant degradation of the carrier image. Also, the quality of the extracted watermarks was determined for different values of the parameters used to regulate the intensity of the attacks. In particular, attacks such as JPG compression, Gaussian noise, Speckle noise and Salt & Pepper are considered. In order to effectively display the large number of obtained results, they are displayed in 3D graphics for all carrier images and all analyzed types of attacks.

In the second chapter, the characteristics of binary images and QR code, which are used as watermarks in this paper, are given. The YCrCb transformation is briefly described, which will enable inserting the watermark into the chrominance component of the carrier image. The third section describes in detail the recommended algorithm for inserting, that is, extracting a watermark from an protected image. In the fourth chapter, the proposed algorithm was evaluated on several color images and several characteristic non-geometric attacks. The obtained results of the performed experiment are shown in appropriate graphs. The results were analyzed based on the objective quality parameters of the extracted watermark. In the fifth chapter, appropriate conclusions are drawn on the application of the proposed concept based on the conducted tests.

II. THEORETICAL BACKGROUND

A. Binary Image

In this work, the term binary image means an image consisting of pixels that can take one of two values, black or white. This allows the value of each pixel to be stored as a single bit, that is, it can have a value of 0 or 1. Binary images take up little memory space, so a 256×256 pixel image only needs to provide 8 kB of memory space. Binary images were previously used to transmit images in fax machines, while today they are used in business automation in barcodes. This paper uses a binary image obtained by encoding some textual information. As already mentioned, the information in this work refers to the owner, the identity of the person depicted in the supporting image or the source of the information. A two-dimensional QR code was used to encode this information.

B. QR code

QR (Quick Response) code is a type of two-dimensional barcode that was designed and developed by Toyota's subsidiary in 1994. Toyota used this code to track vehicles in the manufacturing process, while today the QR code is used in many areas. With the advent of smartphones, QR codes have gained importance because most of these phones have one or more cameras that can be used to scan the QR code and decode it. The QR code has the form of a two-dimensional matrix whose elements can take only two values: logical zero or logical one. If this matrix is printed on paper or displayed on a screen, it represents a binary image. This binary image is divided into different sections that are important for successful decoding. Depending on the type of data and the required information capacity, there are 40 versions of QR codes. The smallest QR code in version 1 has a module size of 21×21 pixels, and can store only 152 bits. On the other hand, version 40 has a dimension of 177×177 pixels and can store 23648 bits. In addition to data memorization, the QR code allows selection of 4 levels of error correction: low, medium, quartile and high. The QR code consists of the following sections: finder patterns, data area, alignment pattern, timing pattern, cell, quit zone, dark module, and separators. A finder pattern is a set of three pattern blocks located at top-left, top-right, and bottom-right locations. Alignment pattern is used for corrections in case of minor image changes during QR code reading. The data area is the central area of the QR code that contains the encoded data. Separators are used to separate the find pattern from the data area. In addition, the QR code provides an improved security mechanism and has error correction algorithms, which is a significant advantage over ordinary binary images that do not have this mechanism.

C. YCbCr color model

Color images can be modeled with different color systems, and the most common are RGB, CMYK or YUV models [9]. For this paper, the YCrCb color model is interesting, which consists of three components: the luminance component (Y) and two chrominance components (Cr and Cb). The relationship between the RGB and YCrCb components of the color model is established by the expressions:

$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} +0.299 & +0.587 & +0.114 \\ -0.169 & -0.331 & +0.500 \\ +0.500 & -0.419 & -0.081 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} + \begin{bmatrix} 0 \\ 128 \\ 128 \end{bmatrix} \qquad (1)$$

In the YCrCb color model, the Y component represents the illuminance in the form of a monochrome image, so by adding the color components (Cb and Cr) it can be easily transformed into a color image. In this paper, a binary watermark - QR code is inserted into one of the image color channels in the YCrCb model. The idea of inserting a watermark into the color channel is inspired by the HVS (Human Visual System) and results in less noticeable degradation of the protected image than when the watermark is inserted into the Y component of the YCrCb model.

With the YCrCb color model, the selection of the sampling scheme is enabled. In order to reduce the size of the image storage file, a downsampling scheme for color components can be used. This technique is standardly used when encoding video sequences. In this paper, a reduced sampling scheme is used, which is denoted as 4:2:0. This has the effect of assigning one sampled value of the color component to the pixels it is surrounded by. In this way, the size of the image storage file is reduced four times. An objectively reduced image quality cannot be subjectively observed by observation. In previous works, the authors inserted color watermarks by decomposing them into multiple binary images. Also, previous works analyzed the performance of algorithms that connect the user's image with his COVID credentials [3]. The previously described features of QR codes have a built-in error correction mechanism, which enables their successful use even in the event of errors.

Fig. 1. Layouts of original carrier images with references used in the experiment a) Lena b) Mandril c) Pepper d) Airplane e) Fruits

III. Insertion/Extraction Algorithm

Watermark insertion is based on block SVD decomposition of the color component of the carrier image. The SVD coefficients of the color component Cb are modulated as a function of the values of all watermark bits. The inverse procedure is used to extract the inserted bits from which the watermark is composed. The insertion and extraction algorithm is presented in multiple IN and EX steps respectively].

Step **IN**$_1$: Transform the carrier image I from RGB to YCrCb color model.

$$\boldsymbol{YCbCr} = rgb2ycbcr(\boldsymbol{I}) \tag{2}$$

Step **IN**$_2$: Divide the chrominance component **Cr** into non-overlapping blocks $\mathbf{H}_{i,j}$ of size 4×4 pixels.

Step **I**$_3$: Perform SVD decomposition for each block $\mathbf{H}_{i,j}$ from the **Cr** component:

$$\boldsymbol{H}_{i,j} = \boldsymbol{U}_{i,j} \times \boldsymbol{S}_{i,j} \times \boldsymbol{V}_{i,j}^{T} \ . \tag{3}$$

Step **IN**$_4$: Modify the elements from the second and third rows of the first column of each $\mathbf{U}_{i,j}$ matrix (elements $u_{2,1}$ and $u_{3,1}$) based on the value of each individual bit w from the QR code image:

$$if \ \ w = 1, \begin{cases} u_{2,1}^{*} = sign(u_{2,1}) \times \left(U_{avg} + \frac{T}{2}\right) \\ u_{3,1}^{*} = sign(u_{3,1}) \times \left(U_{avg} - \frac{T}{2}\right) \end{cases} \tag{4}$$

$$\text{else } if \ \ w = 0, \begin{cases} u_{2,1}^{*} = sign(u_{2,1}) \times \left(U_{avg} - \frac{T}{2}\right) \\ u_{3,1}^{*} = sign(u_{3,1}) \times \left(U_{avg} + \frac{T}{2}\right) \end{cases} \tag{5}$$

$$U_{avg} = \frac{(|u_{2,1}| + |u_{3,1}|)}{2} \tag{6}$$

where T is the differential threshold and it is parameter of the algorithm used to determine the insertion strength. The modified matrix is denoted by $\mathbf{U}^*_{i,j}$.

Step **IN**$_5$: Perform the inverse SVD transformation of each block to obtain the block with the bit inserted from the QR code.

$$\boldsymbol{H}_{i,j}^{*} = \boldsymbol{U}_{i,j}^{*} \times \boldsymbol{S}_{i,j} \times \boldsymbol{V}_{i,j}^{T} \tag{7}$$

Step N**I**$_6$: Place the block with the inserted QR code bit in the appropriate place in the protected carrier photo of the person. Repeat steps **I**$_4$ and **I**$_5$ for all bits from the QR code.

$$\boldsymbol{RGB}^{*} = ycbcr2rgb(\boldsymbol{I}^{*}) \tag{8}$$

In this way, all bits from the QR code are inserted into the carrier image. In order to extract the watermark inserted in this way, it is not necessary to have the originals of either the image or the watermark, so this algorithm belongs to the class of BLIND algorithms. The algorithm for extracting a watermark from a protected image is shown in a series of EX steps.

Step **EX**$_1$: Perform transformation of protected carrier image A' from RGB to YCrCb color model.

$$\boldsymbol{YCbCr} = rgb2ycbcr(\boldsymbol{I}') \tag{9}$$

Step **EX**$_2$: Divide the **Cb** component of the protected image into non-overlapping blocks $\mathbf{H'}_{i,j}$ of dimensions 4×4 pixels.

Step **EX**$_3$: Perform SVD decomposition over all blocks $\mathbf{H'}_{i,j}$ carrying image.

$$\boldsymbol{H}_{i,j}' = \boldsymbol{U}_{i,j}' \times \boldsymbol{S}_{i,j}' \times \boldsymbol{V}_{i,j}'^{T} \tag{10}$$

Step **E**$_4$: The value of the corresponding extracted watermark bit w' is obtained by applying the following expressions:

$$w' = 1, \begin{cases} 0, \ if \ u_{2,1}' > u_{3,1}' \\ 1, \ if \ u_{2,1}' \leq u_{3,1}' \end{cases} \tag{11}$$

Step **EX**$_5$: Set the extracted bit value to the appropriate location in the watermark.

Step **EX**$_6$: Repeat steps **E**$_4$ and **E**$_5$ for all blocks of the protected carrier photo and form the complete binary image of the QR code. The resulting binary image represents the extracted QR code.

IV. Eksperimentalni Rezultati

In the experimental part of this work, the well-known color images of Lena, Mandrill, Pepper, Airplane and Fruits were used in a resolution of 512×512 pixels. Layouts of these images, which were used as carrier images, are shown in Fig. 1. Below each carrier image is given their reference used in the experimental part of the work. A watermark is inserted into all carrier images using the proposed algorithm. As already explained, a watermark obtained by encoding the text "Lena Forsen (Sjooblom), Swedish model" with a QR code was used. In this way, a binary image with dimensions of 64×64 pixels was created. The layouts of the original Leno carrier image and the original watermark are shown in Fig. 2a.

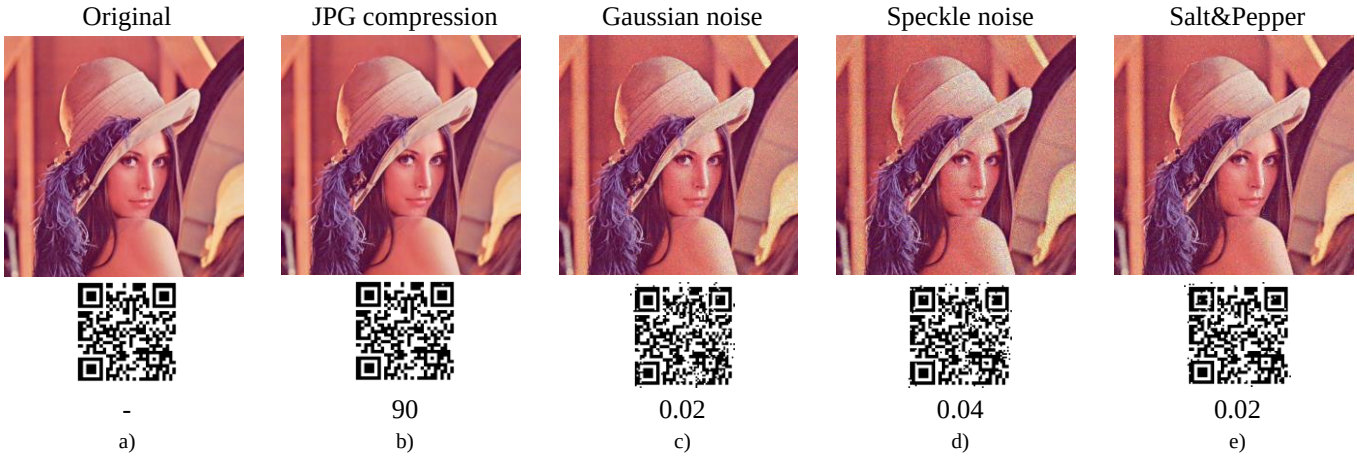


Fig. 2. Layouts a) unprotected carrier image "Lena" and original watermark b) protected image "Lena" with attack in the form of JPG compression defined quality 90 and extracted watermark c) protected image "Lena" with attack in the form of Gaussian noise with variance 0.02 and extracted watermark d) protected image of "Lena" with attack in the form of Specle noise with variance 0.04 and extracted watermark d) protected image of "Lena" with attack in the form of Salt &Pepper and noise density 0.02 and extracted watermark.

The insertion algorithm allows selection of the insertion strength, so that for each carrier image a series of protected images with different values of the differential threshold T is created. A higher value of the differential threshold T will cause a greater degradation of the protected image, but also the extraction of a better quality watermark. The proposed algorithm defines that the watermark is inserted only in the Cr chrominance channel of the supporting image. In the experiment, 6 protected images with different differential insertion threshold T=[0.1 0.15 0.2 0.25 0.3 0.35] were generated for all supporting images. In total, 30 different protected images were generated, on which different attacks are applied. Watermarks of different quality are extracted from such attacked protected images.

In previous works, the problem of choosing the differential threshold T was considered and its optimal value was determined without the presence of attacks. In the conditions of attacks on protected images, it is clear that a higher value of the differential threshold T must be chosen in order to ensure high quality of the watermark. In the experimental part of the work, non-geometics attacks were analyzed, which include JPG compression, Gaussian, Speckle noise and Salt & Pepper noise. The quality of the extracted watermarks is evaluated on the basis of objective parameters based on the absolute number of bit errors ERROR and normalized correlation coefficients NCC in relation to the original watermark. The number of errors in the extracted watermark, which is considered a bit image, is determined according to the expression (12).

$$ERROR(w,w') = \sum_{k=1}^{K}\sum_{l=1}^{L}[w(k,l) \otimes w'(k,l)] \quad (12)$$

where $w(k,l)$ refers to the original and $w'(k,l)$ refers to the extracted watermark. The variables K and L are the dimensions of the watermark, while $\otimes$ is the XOR operator. The normalized correlation coefficient NC is determined according to expression (13).

$$NC(w,w') = \frac{\sum_{k=1}^{K}\sum_{l=1}^{L}[w(k,l)\times w\prime(k,l)]}{\sqrt{\sum_{k=1}^{K}\sum_{l=1}^{L}w^2(k,l)}\times\sqrt{\sum_{k=1}^{K}\sum_{l=1}^{L}w^2(k,l)}} \quad (13)$$

where $w(k,l)$ refers to the original and $w'(k,l)$ refers to the extracted watermark. The variables K and L are the dimensions of the watermark.

The results of the experiment are shown by a series of graphs in Fig. 3 related to non-geometrics attacks: JPG Compression, Gaussian noise, Speckle noise and Salt&Pepper noise. The 3D plots in these images show the quality of the extracted watermarks based on the absolute number of errors in the extracted watermarks. From Fig. 3, it can be seen that the number of errors in the extracted binary watermark increases with the decrease of the differential threshold, which was expected. Also, the number of errors increases with the intensity of attacks. With JPG compression, as the JPG quality parameter increases, the number of errors decreases and vice versa. The graph clearly shows the parameter space in which there are no bit errors in the extracted watermarks. It can be seen that the forms of the graphics are similar, but that the details can differ significantly - all depending on the supporting image. In Fig. 2b) to 2e) shows the layout of the protected image of Lena for all considered attacks: JPG Compression, Gaussian noise, Speckle noise and Salt&Pepper. The intensity of individual attacks is shown below each of the extracted trademarks. Also below the protected image are shown the layouts of the extracted watermarks. Bit errors in the extracted watermarks can be observed on them. These errors can make it difficult to decode QR codes and will be the subject of further research. In Fig. 2 is shown only smaller than the results obtained in the experiment.

In Fig. 4 shows an example of a 3D graphic related to the normalized correlation coefficients NC as a function of the differential threshold and the intensity of the applied attack. The selected graphics can serve as an illustration of the obtained results and refer to the proprietary Lena carrier image subjected to various intensities of JPG compression, Gaussian noise, Speckle noise and Salt &Pepper. From all the graphics in Fig. 4, it can be seen that the NC values tend to unity with the increase in the differential threshold and the decrease in the intensity of individual attacks. Other graphs related to the values of NCC coefficients determined for other carrier images and attack types are not shown due to lack of space.

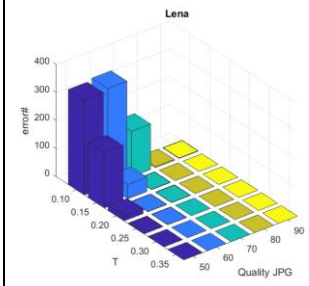
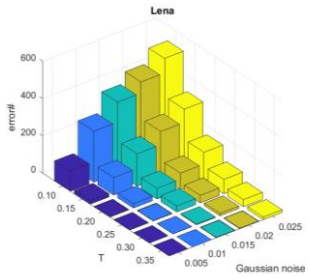
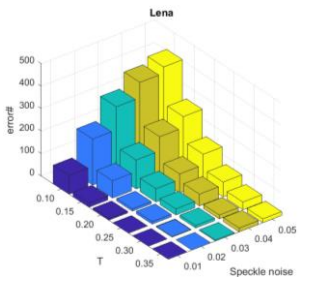
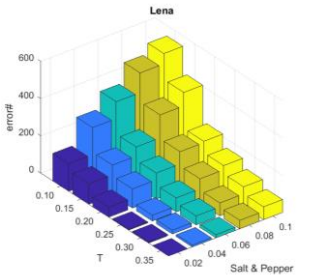
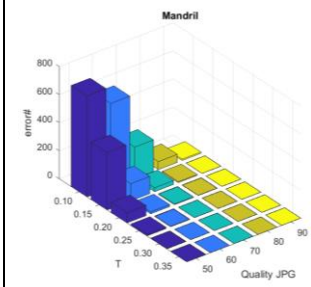
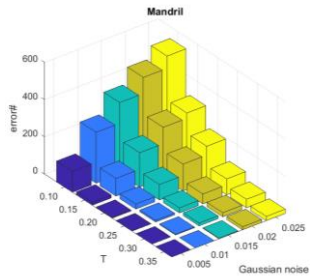
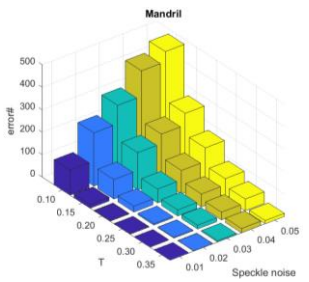
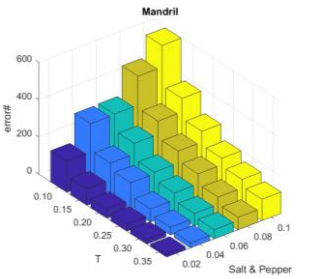
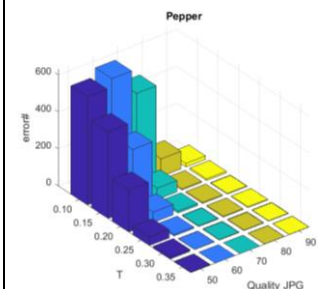
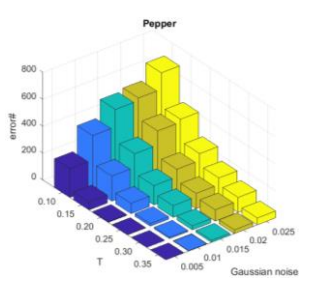
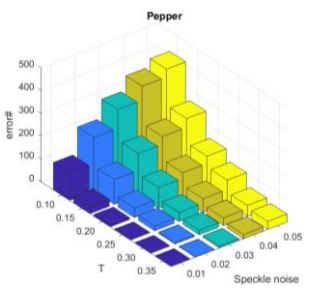
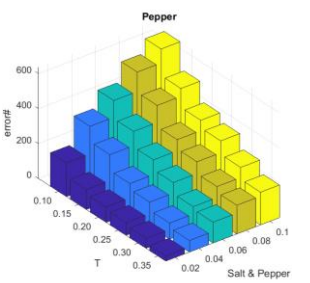
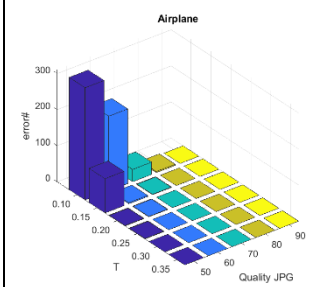
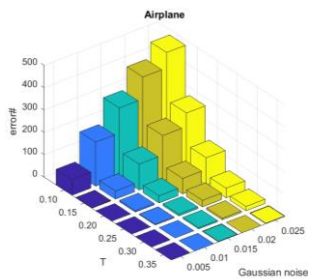
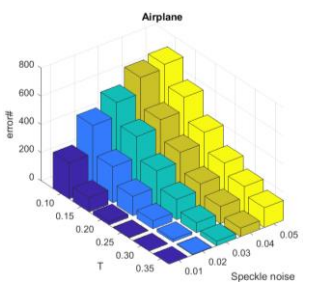
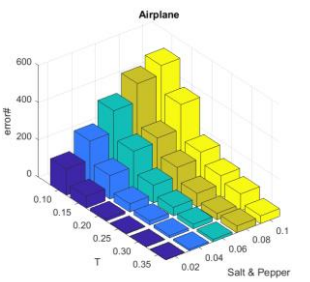
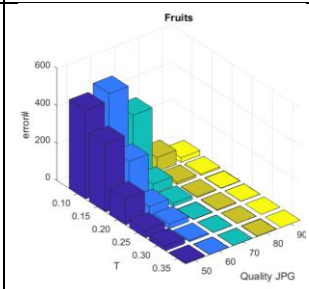
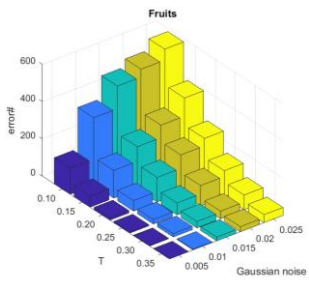
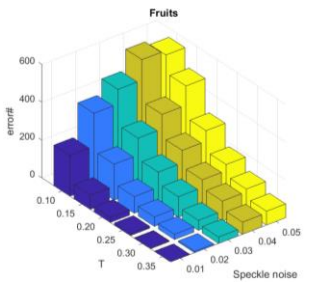
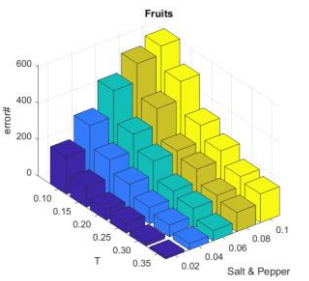
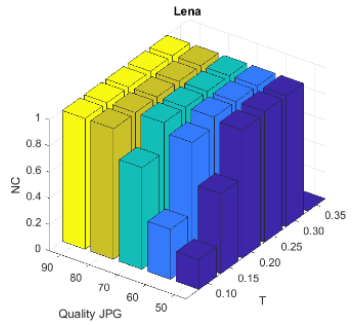
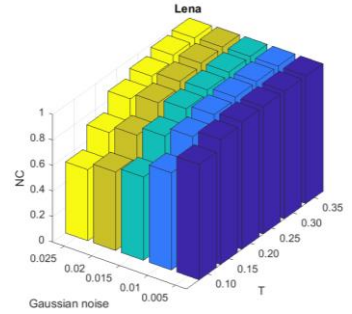
#	ATTACKS Number of errors			
	JPG Compression	Gaussian noise	Speckle noise	Salt & Pepper noise
1				
2				
3				
4				
5				
	a)	b)	c)	d)

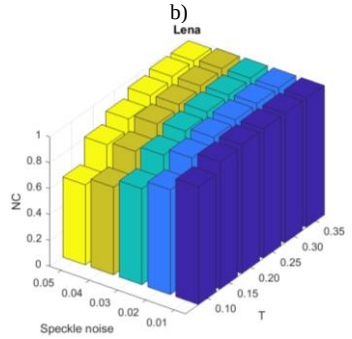
Fig. 3. Number of errors in the extracted watermark for all carrier images as a function of differential threshold and attack intensity a) JPG Compression b) Gaussian noise c) Speckle noise d) Salt & Pepper noise.



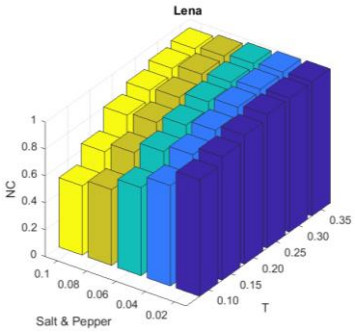
a)



b)



c)



d)

Fig. 4. The value of NCC coefficients for the carrier image of Lena for a) JPG b) Gaussian noise c) Speckle noise and d) Salt&Pepper attack.

All displayed graphs have a similar shape, and the values of the NC coefficients tend to unity in a similar way, as stated previously. The presented algorithm enables the desired performance by selecting the appropriate differential threshold. A higher value of the differential threshold can provide greater resistance to attempts to remove the watermark from the protected image. From the attached graphics in Fig. 3 and Fig. 4, it can be concluded that the performance depends on the supporting image, but that they do not have a significant influence on the behavior of the algorithm itself. The degradation of carrier images caused by attacks is far more significant than the degradation caused by watermark insertion alone. This has the effect that the watermark inserted by the recommended algorithm can be extracted even in conditions of large attacks. The recommended algorithm can provide a compromise between insertion strength, carrier image degradation, desired quality of the extracted watermark, and attack intensity.

V. Conclusion

In this paper, the performance of the recommended image protection algorithm against copying and illegal use was analyzed. A binary watermark in the form of a QR code obtained by encoding the desired information is inserted into the chrominance Cr component of the carrier image. The performance of the recommended algorithm when trying to remove the watermark from the image was analyzed. The resistance of the recommended algorithm to non-geometrics attacks, which are standardly used for watermark removal, is discussed. By varying the strength of insertion and the intensity of the attacks, it was established that it is possible to extract the watermark without errors for all types of considered attacks. By decoding the extracted QR code, information about the source or owner of the image can be obtained. In further research, the additional ability of QR codes to correct possible errors in the extracted watermark will be considered, so it can be influenced by lowering the value of the differential threshold in the insertion process. The results of the experiment confirm that the presented concept can be successfully used in copyright protection, that is, in the protection of images from copying and illegal use and distribution.

References

[1] Cisco Annual Internet Report (2018–2023), White paper

[2] C. Neubauer, J. Herre, “Advanced Watermarking and its Applications”, Fraunhofer Institute for Integrated Circuits IIS, https://www.iis.fraunhofer.de, 30.01.2023.

[3] Z. Veličković, S. Veličković, Z. Milivojević, „Application of Watermark in the Form of QR Code in COVID Certificate Validation“, Journal of Mechatronics, Automation and Identification Technology JMAIT, Vol. 6, No. 2, pp. 1 – 5, 2021.

[4] S. B. B. Ahmadi, G. Zhang, M. Rabbani, L. Boukela and H. Jelodar, “An intelligent and blind dual color image watermarking for authentication and copyright protection”, Applied Intelligence, Springer Science+Bussiness Media, 2020.

[5] A. M. Cheema, S. M. Adnan, Z. Mehmood, “A novel Optimized Semi-Blind Scheme for Color Image Watermarking”, IEEE Acces, Vol. 8, pp. 3169525-169547, 2020

[6] C. Patvardhan, P. Kumar, C. V. Lakshmi, “Effective color image watermarking scheme using YCbCr color space and QR code”, Multimed. Tools Appl. (2018) 77:12655–12677.

[7] B. Wang, P. Zhao, “An Adaptive Image Watermarking Method Comining SVD and Wang-Landau Sampling in DWT Domain,”, Mathematics, 2020, 8, 691, doi:10.3390/math80506

[8] P. P. Thulasidharan, M. S. Nair, “QR code based blind digital image watermarking with attack detection code”, Int. J. Electron. Commun. (AEÜ) 69 (2015) 1074–1084.

[9] I. E. Richardson, The H.264 Advanced Video Compression Standard, John Wiley & Son Ltd, 2010.