symmetry

MDPI

Article

EasyStego: Robust Steganography Based on Quick-Response Barcodes for Crossing Domains

Zhenhao Luo *, Wei Xie *, Baosheng Wang, Yong Tang and Qianqian Xing

College of Computer, National University of Defense Technology, Changsha 410073, China; bswang@nudt.edu.cn (B.W.); ytang@nudt.edu.cn (Y.T.); xingqianqian12@nudt.edu.cn (Q.X.)

* Correspondence: luozhenhao13@nudt.edu.cn (Z.L.); xiewei@nudt.edu.cn (W.X.)

Received: 22 January 2019; Accepted: 12 February 2019; Published: 14 February 2019

Abstract: Despite greater attention being paid to sensitive-information leakage in the cyberdomain, the sensitive-information problem of the physical domain remains neglected. Anonymous users can easily access the sensitive information of other users, such as transaction information, health status, and addresses, without any advanced technologies. Ideally, secret messages should be protected not only in the cyberdomain but also in the complex physical domain. However, popular steganography schemes only work in the traditional cyberdomain and are useless when physical distortions of messages are unavoidable. This paper first defines the concept of cross-domain steganography, and then proposes EasyStego, a novel cross-domain steganography scheme. EasyStego is based on the use of QR barcodes as carriers; therefore, it is robust to physical distortions in the complex physical domain. Moreover, EasyStego has a large capacity for embeddable secrets and strong scalability in various scenarios. EasyStego uses an AES encryption algorithm to control the permissions of secret messages, which is more effective in reducing the possibility of sensitive-information leakage. Experiments show that EasyStego has perfect robustness and good efficiency. Compared with the best current steganography scheme based on barcodes, EasyStego has greater steganographic capacity and less impact on barcode data. In robustness tests, EasyStego successfully extracts secret messages at different angles and distances. In the case of adding natural textures and importing quantitative error bits, other related steganography techniques fail, whereas EasyStego can extract secret messages with a success rate of nearly 100%.

Keywords: steganography; cross-domain; QR barcodes; confidentiality; physical distortions

1. Introduction

Steganography is a technique of concealing secret data within a carrier file that aims to protect the confidentiality of the data. With the gradual emergence of user-data collection as a trend, user shopping information, social information, and search information are increasingly collected and utilized by interested organizations. However, a series of information-leakage incidents (for example, Quora [1], Facebook [2], MyFitnessPal [3]) have led users to pay more attention to their personal information and, in turn, have increased research attention on steganography. Some researchers have used digital steganography to protect the confidentiality of patient information [4–6] and to prevent the detection of secret conversations between users. These steganography schemes can protect secret data from being leaked or detected when passed on in the cyberdomain.

However, leaks of user information are not confined to the Internet. Many information leaks come from express packages, banks, and hospitals [7]. Information printed on paper in plain text is easily accessible by anyone without any technology. For example, a cleaner at a bank can easily get users' transaction information through waste papers in the trash, and any staff member in a hospital can acquire the basic information and conditions of patients from labels in the ward and illegally sell them

to other institutions. Therefore, user information should be protected in the physical domain as well as the cyberdomain. With the advent of quick-response (QR) barcodes [8], it is possible to use digital technology to protect user information in the physical domain using these barcodes as carriers. We can encrypt a user's sensitive information using an encryption algorithm and hide it in a QR barcode without affecting the normal operation of the barcode. We find that the intrinsic characteristics of information hiding have implicit symmetry in the physical domain.

To better visualize how the scheme can be utilized, the following scenario was devised. A user wants to mail an express package. They encrypt the related information of the sender and the recipient and hides this information in a QR barcode. Others can acquire only the normal information in the QR barcode (e.g., the delivery address) but cannot access sensitive information (for example, names, phone numbers and other package content). Only the actual delivery courier can scan the relevant information necessary for mailing. For all of these operations, the sender needs to print only the QR barcode on the application. Due to the wide distribution and large amount of information, these situations require that the scheme be uncomplicated in its operation and capable of working under various physical conditions (e.g., light conditions, distances, angles) and that the hidden information be robust to physical distortions introduced during the printing process. Some existing digital steganography schemes [4,9–12] do not work properly in this scenario because the hidden data are destroyed. Other steganography schemes [13–15] based on QR barcodes are not easy to use in the real world because they are not robust to complex physical factors.

In this paper, we define the concept of cross-domain steganography and propose EasyStego, a novel cross-domain and antiphysical distortion steganography scheme. EasyStego is based on the use of QR barcodes as carriers; therefore, it is robust to the physical distortions of the physical domain. Based on the error tolerance of the QR barcode, EasyStego can conceal a large number of secret messages without further increasing the size of the carrier file itself. EasyStego uses the arrangement of QR barcodes to optimize the embedding strategy so that the embedded secret payload destroys as little of the QR data as possible. EasyStego is the first image steganography scheme suitable for use in various complex environments. The experiments in Section 5 specifically illustrate the unique advantages of EasyStego and the differences between EasyStego and other steganography schemes.

EasyStego achieves the above goals as follows: (i) it implements a novel image steganography scheme between different physical environments that can resist physical distortions and cross domains while carrying secret messages; (ii) it has a large capacity for secret messages of up to 16420 bits, and its secret payload has little impact on the extraction of the QR data; and (iii) it has strong robustness and can extract secret messages properly under complex physical conditions in which other related steganography schemes fail.

The contributions of this paper are as follows:

- We propose a new type of steganography, namely, cross-domain steganography, and present its specific definition for the first time. In contrast to traditional digital steganography, cross-domain steganography can be transferred and decoded in the physical domain and cyberdomain, and it can exist in the physical domain as an object.
- We propose and implement EasyStego, a novel antiphysical distortion and cross-domain steganography scheme for complex physical environments. EasyStego is based on the use of QR barcodes as carriers; therefore, it is robust to physical distortions of the physical domain. Additionally, EasyStego has large embeddable secret capacity, strong robustness, and good concealment, and can work normally in a variety of situations.
- We carefully evaluate EasyStego in a real physical environment with different complex physical factors. Evaluation in real-world scenarios demonstrates the effectiveness of our steganography scheme.

The main novelty of EasyStego is its ability to cross domains, and the novel integration of techniques (some of which already exist) that provides strong robustness with minimal adverse

impacts. EasyStego, which is based on the use of QR barcodes as carriers, has strong robustness to work in complex physical environments. A large extensive set of real-world experiments empirically prove its efficacy and verify its viability.

The remainder of this paper is organized as follows. In Section 2, we describe relevant works in related fields and compare them with EasyStego. In Section 3, we describe usage scenarios and define the concept of cross-domain steganography. In Section 4, we thoroughly describe the steganography scheme of EasyStego. Section 5 reports the evaluation results and compares EasyStego with existing steganography schemes. Section 6 shows the results of an application scenario simulation in the real world. In Section 7, we discuss the reasons why EasyStego performs well. Finally, Section 8 concludes the paper.

2. Related Work

Steganography is the practice and technique of concealing secret messages within a carrier file. By concealing the existence of secret messages, steganography prevents their detection by others. Steganography can generally be divided into traditional and modern steganography.

Traditional steganography uses physical, chemical, or other methods to conceal secret messages in the physical domain and prevents the existence of secret messages from being perceived by anyone other than the intended receiver. The first recorded use of steganography can be traced back to 440 BC. In this case, a Greek ruler sent a secret message to his subordinate by shaving the head of a servant, writing the secret message onto his scalp, and then sending him to the destination once his hair had regrown [16]. After continuous development, various types of steganography schemes using physical and chemical methods appeared. For example, a secret message written with colorless ferric sulfate becomes bright blue after wiping with potassium cyanate. Due to the complexity of the conditions and operations required for decryption, it is difficult for such schemes to be processed by a computer. Therefore, it is difficult for traditional steganography to cross from the physical domain to the cyberdomain.

Modern steganography, also known as digital steganography, conceals secret messages in digital carrier files. Currently, carrier-file types include image, audio, and multimedia files. Since EasyStego is a cross-domain image-steganography scheme, the following part of this section mainly introduces image steganography. Interested readers should refer to References [17–19] for more information on audio and video steganography.

Mehboob et al. [20] proposed a steganography scheme to conceal secret messages in a colorful image in BMP format using the least significant bit (LSB). In the LSB method, the eighth bit of every carrier-file byte is substituted by one bit of secret information. To maintain high camouflage image quality with high hiding capacity, Lu et al. [21] proposed a novel approach based on the LSB matching method for embedding using a dual-imaging technique, and its camouflage images are also of an above-average level and more difficult to detect by others.

Bansod et al. [22] proposed a steganography scheme using hybrid cryptography. This method is based on the use of the DES and RSA algorithms to achieve secret-message encryption, which increases the difficulty of cracking secret messages. Ramesh et al. [23] proposed an image-steganography scheme based on frequency-domain processing. It uses discrete-wavelet transform (DWT) to conceal secret messages in the carrier image as a secret image, and then sends the secret image containing the secret messages together with the original carrier file. After receiving the file, the secret message is obtained by subtracting the secret-image DWT component from the original cover-image DWT component. Islam [24] used DWT-SVD to conceal a QR barcode in a carrier image with 100% message recovery. Liao et al. [4] proposed an adaptive steganography scheme that preserves the dependencies of the interblock discrete-cosine transform (DCT) coefficients. It can be applied to medical JPEG images, and has better antisteganographic detection performance. Denemark et al. [10] embedded a secret message in the quantized DCT coefficients of an image in JPEG format by adding independent realizations of the heteroscedastic noise. Their experiments showed that a large number of payloads could be

embedded in monochrome sensors or low-quality JPEG images.

Roshan Shetty et al. [25] proposed an interesting scheme that hides secret messages using Sudoku puzzles. They used a 27 by 27 reference matrix to conceal secret messages using Sudoku, and selected more suitable regions in the reference matrix to minimize the distortions. Tkachenko et al. [14] proposed a steganography scheme called 2LQR based on maximizing the correlation values between P and S degraded patterns and reference patterns. This scheme carried secret messages by increasing the texture pattern of the QR barcode.

However, the above related technologies and methods are difficult to apply across domains because their data cannot be resistant to physical distortion.

Some steganography schemes use the error-correction capability of QR barcodes to hide secret payloads. Due to the inherent features of QR barcodes, they are not easily noticed and are resistant to physical distortion. Lin et al. [13] observed and proposed a novel scheme to conceal secret messages in QR barcodes. This method uses the error-correction capability of QR barcodes to hide secret messages without distorting the readability of the barcodes. Its payload arrangement is adjustable according to the selection of the QR version and the error-correction level. Simulations demonstrated that the scheme has low computational complexity but no error-correction ability; when any bit of the carrier QR barcode is damaged, it is difficult to recover the secret payload. Lin et al. [15] improved the steganography scheme of Reference [13] and used the concept of the exploiting modification direction (EMD) scheme to hide secret payloads in QR barcodes. Compared with Reference [13], Reference [15] improved the hiding capacity, and the secret payload could be recovered with few errors. In the experimental part, we mainly compare our work with References [13,15].

EasyStego, the novel steganography scheme proposed in this paper, is more resistant to physical distortion than the above steganography schemes and can cross domains (from the cyberdomain to the physical domain, and from the physical domain to the cyberdomain). A comparison of the methods is shown in Table 1. EasyStego uses the error-correction capability of QR barcodes to hide data and uses the arrangement of QR barcodes to optimize the hidden location of the secret payload. Due to the widespread use of QR barcodes, EasyStego carriers would not easily be noticed by adversaries. Additionally, EasyStego uses Reed–Solomon codes for error correction, such that encoded secret messages can be extracted even if some bits are broken.

Table 1. Comparison of EasyStego with popular steganography schemes.

Scheme	Embedding Method	Encryption	Carrier	Need Original File	Error Correction	Resist Physical Distortion	Impact on the Original Image	Cross-Domain
[20]	LSB	No	BMP image	No	No	×	Low	×
[22]	LSB	DES & RSA	image	No	No	×	Low	×
[23]	DWT	No	image	Yes	No	×	Low	×
[24]	DWT-SVD	No	image	No	No	×	Low	×
[4]	DCT	No	JPEG image	No	No	×	Low	×
[10]	DCT	No	JPEG image	No	No	×	Low	×
[14]	Texture pattern	No	QR barcode	No	No	×	High	×
[25]	-	No	Sudoku (uncommon)	No	No	✓	-	✓
[13]	Error correction capability	-	QR barcode	No	No	✓	High	✓
[15]	EMD	-	QR barcode	No	Yes	✓	High	✓
EasyStego	Error correction capability and arrangement	AES	QR barcode	No	Yes	✓	Low	✓

3. Concept of Cross-Domain Steganography

3.1. Usage Scenarios

As the trend of collecting user data has gradually emerged, user-information leakage occurs not only on the Internet but also in the physical domain.

Hospital information, bank information, and express-package information are frequently stolen by anonymous parties. Unlike the cyberdomain, stealing information in the physical domain does not require advanced technology. For example, a cleaner at a bank can easily obtain users' transaction information from waste papers in the trash, and any hospital staff member can acquire the basic information and conditions of patients from labels in the ward and illegally sell this information to other institutions. Therefore, a scheme is required to protect sensitive user information in the physical domain. To better visualize how the scheme can be utilized, the following scenarios are devised.

Alice wants to mail an express package to Bob. She encrypts the sensitive information of the sender and the recipient and hides them in a QR barcode. Others can only obtain the normal information of the QR barcode (e.g., the delivery address) and not the sensitive information (for example, names, phone numbers, and other package content). Only the actual delivery courier can scan the relevant information necessary for mailing. In this way, malicious postmen cannot obtain more sensitive information about Alice and Bob from the express order because they do not have the corresponding permissions. When Bob signs for the package, he can only decrypt the QR barcode with the key that Alice shared with him. Alice can easily perform all of these operations with a tap on a smartphone.

In another scenario, Bob develops a disease and is hospitalized. The label on the side of the bed is convenient for the attending doctor to understand his condition and some basic information, but Bob does not want other people who are not relevant to know his situation. At this point, a scheme is needed that makes his sensitive information inaccessible to unrelated people, while in an emergency, relevant people can quickly obtain this information. In these scenarios, the subjects are eager for a steganography scheme that protects the confidentiality of their sensitive information in the physical domain.

3.2. Concept of Cross-Domain Steganography

According to the above usage scenarios, we can define cross-domain steganography. Cross-domain steganography is the practice of using carriers that can be extracted in various domains to conceal a file, message, image, or video. The cross-domain in this article refers to the cross-domain between the cyberdomain and the physical domain. The cyberdomain refers to the binary world composed of digital data, and the physical domain refers to the natural world in which people live.

3.3. Advantages and Challenges

Cross-domain steganography has an advantage that current steganography does not have: it is resistant to physical distortion and can recover a secret message despite extensive noise. These features provide flexibility across domains.

However, cross-domain steganography presents higher challenges on steganography schemes. It requires:

- steganographic carriers that are present in the physical domain;
- steganographic carriers that are not too special;
- steganography schemes that can resist the adverse effects of physical distortion (e.g., angles and distances); and
- steganography schemes with error correction to ensure the readability of secret messages under complex physical conditions (e.g., abrasions and wrinkles).

4. Proposed EasyStego Scheme

4.1. Motivation

The currently popular steganography schemes cannot easily satisfy all cross-domain requirements.

Traditional steganography schemes cannot meet cross-domain requirements. For example, a secret message written with colorless ferric sulfate becomes bright blue when wiped with potassium cyanate. Due to the complexity of the conditions and operations required for decryption, it is difficult for this scheme to be processed by a computer.

Some works [4,10,20,22–24] may have performed well in the cyberdomain, but their secret messages are difficult to extract in the physical domain in the case of physical distortion. For example, when an image is displayed or printed, the image may not actually be able to be displayed due to the device's own factors (for example, display resolution and rendering, and printer printing effect). Additionally, when the scanning device extracts a secret message, the angles, distances, focus, light, and the device itself can cause the physical distortion of the image. Due to these factors, these approaches are unable to cross domains. Tkachenko's work [14] is better than the above schemes, but the extraction of the secret message is still difficult because of the extremely small texture pattern. Although the Sudoku puzzle strategy [25] can work in the physical domain, it is not practical, because Sudoku puzzles are not widely used, and they can easily attract the attention of adversaries.

References [13,15], which are based on the error correction of QR barcodes, resist the physical distortions caused by angles and distances. Their carriers are popular in the real world and can be present in the physical domain. However, when physical conditions become increasingly more complicated (e.g., abrasions and wrinkles), they also encounter difficulty in working properly, making it difficult to recover secret payloads in complex physical environments.

To work in complex physical environments, we propose a novel steganography scheme, EasyStego, in this paper. It is based on the use of QR barcodes as carriers, is more resistant to physical distortion than the above steganography schemes, and can cross domains. Due to the widespread use of QR barcodes, EasyStego carriers would not easily be noticed by adversaries.

4.2. Overview

EasyStego works properly in complex physical environments and maintains high hiding capacity. To achieve these goals, we implemented EasyStego as follows.

EasyStego is based on the use of QR barcodes as carriers; therefore, it is robust to physical distortions that can occur in complex physical environments. People cannot read the content of QR barcodes directly from the semantics without the use of devices (e.g., smartphones and scanners). If a secret message is embedded without changing the content of the QR barcode, users generally do not notice the message. Additionally, QR barcodes are widely used in various settings. EasyStego using QR barcodes as carriers can thus effectively reduce unwanted attention.

To embed more secret messages, EasyStego uses dynamic Huffman coding to recode secret messages. It can compress different secret messages using different dictionaries according to the frequency of characters.

To overcome the effects of complex physical environments, EasyStego uses Reed–Solomon codes for the error correction of the embedded data. The use of these codes allows the EasyStego secret message to be normally extracted even if some bits are broken. In addition, EasyStego includes an efficient encoding algorithm that increases the frequency of "\x00" in the encoded bytes, which reduces the difference between the secret image and the original image while improving the capacity of carrying secret messages.

4.3. Carriers

The carriers of the EasyStego scheme are QR barcodes. The QR code was invented in 1994 by Denso Wave [8], and was approved as an international standard (ISO / IEC18004) in June 2000 [26].

According to the international standard [26], QR barcodes provide a total of 40 different versions of data capacity. Version 1 is 21 × 21 modules (the module is the smallest unit in a QR barcode), and each additional version adds four modules of length and width. The largest version, 40, is a 177 × 177 module that can hold up to 7089 numeric characters, 4296 alphabetic characters, or 2953 8-bit binary numbers. QR barcodes have high error correction capability. Each version has four levels of error correction: L, M, Q, and H. Table 2 lists the error-correction levels (ECL) of QR barcodes. Table 2 shows that some QR barcodes with 30% damage can still be read, indicating strong resistance to damage.

Table 2. Error-correction levels of QR barcodes.

Error-Correction Level	Recovery Capacity (%)
Low (*L*)	7
Medium (*M*)	15
Quartile (*Q*)	25
High (*H*)	30

Recovery capacities in this table are approximations.

Table 3 shows the details of the different versions under error-correction level H. In Table 3, the maximum capacities of versions 1, 10, 20, 30, and 40 with error-correction level H are listed. QR data with larger capacity are generally divided into several data groups according to the QR version. For example, the group number is 1 for QR version 1-H, and QR version 10-H has eight groups. It can be seen that the higher the QR version and error-correction level are, the larger the number of QR data blocks. Error-correction codes corresponding to each QR data group are also individually generated.

Table 3. Maximum capacities of QR barcodes (ECL:H).

Version	Maximum Capacity (bits)	Blocks	Error-Correction Code	Maximum Error Pixels
1	72	9 × 1	17 × 1	8 × 8
10	976	15 × 6 + 16 × 2	28 × 8	112 × 8
20	3080	5 × 15 + 16 × 10	28 × 25	325 × 8
30	5960	15 × 23 + 16 × 25	30 × 48	720 × 8
40	10,208	15 × 20 + 16 × 61	30 × 81	1215 × 8

For example, Figure 1 shows the groups in a QR barcode (version 7, ECL:H). As shown in Figure 1, each color represents a group, and there are five groups in total. Additionally, there are 39 blocks in each group, except for the blue group, which has 40 blocks. Each of these groups consists of data codewords and corresponding error-correction codes. Since error-correction codes are separately generated, the error correction of the QR barcode is also performed in groups. Once one of the groups exceeds the threshold of error pixels, the entire QR barcode cannot be read. Based on the arrangement of QR barcodes, EasyStego optimizes the hidden location of the secret payload to reduce the impact on the original QR barcode.

Figure 1. Groups in a QR barcode.

EasyStego selects QR barcodes as carriers for four reasons. QR barcodes (i) are widely used and not easily noticed by adversaries; (ii) are present in the physical domain and are robust to physical distortions; (iii) have high error-correction capability; and (iv) have no semantics for people, and people cannot directly read the content of QR barcodes. It is generally difficult for users to detect modified pixels of the QR barcode, as long as the read content does not change.

4.4. Huffman Coding

Huffman coding [27] is an optimal prefix code that is generally used for lossless data compression. It chooses the representation of each symbol based on the weight of each symbol. Let us consider a dataset $A = (a_1, a_2, \cdots, a_n)$ to be compressed. Then, given a tuple $W = (w_1, w_2, \cdots, w_n)$, where w_i is the weight of a_i, $0 \leqslant i \leqslant n$. Then, calculate $R(A, W) = (r_1, r_2, \cdots, r_n)$, where r_i is the codeword for a_i. Minimize $L(R(A, W))$:

$$L\left(R\left(A, W\right)\right) = \sum_{i=1}^{n} w_i \times \text{length}\left(c_i\right) \tag{1}$$

By optimizing $R(A, W)$, $L(R(A, W))$ is minimized. Finally, an optimized representation set R of A is obtained.

4.5. Reed–Solomon Codes

Reed–Solomon codes [28] are a group of error-correcting codes. Reed–Solomon codes map a data vector $x = (x_1, \ldots, x_k) \in F^k$ into the polynomial p_x with

$$p_x(a) = \sum_{i=1}^{k} x_i a^{i-1} \tag{2}$$

In Equation (2), a is a nonzero element of F^k. p_x is evaluated at n different elements $a_1, \ldots, a_n$ of the field F. The Reed–Solomon codes of data x can be calculated as follows:

$$C(x) = x \cdot \begin{bmatrix} 1 & \ldots & 1 & \ldots & 1 \\ a_1 & \ldots & a_k & \ldots & a_n \\ a_1^2 & \ldots & a_k^2 & \ldots & a_n^2 \\ \vdots & \ldots & \vdots & \ldots & \vdots \\ a_1^{k-1} & \ldots & a_k^{k-1} & \ldots & a_n^{k-1} \end{bmatrix} \tag{3}$$

With Equation (3), data x can be mapped into $C(x)$, which has a dimension higher than that of x. Therefore, transmitted data x have extra information that can help us recover the original data of dimension k if there exist errors in transmitted data x.

4.6. Embedding Procedure

The embedding procedure of EasyStego can be divided into seven steps. The steps are listed below:

Step 1. Get the data arrangement of carrier QR barcode Q and ecc, which is the number of error-correction codes, and calculate the maximum capacity of the secret payload for QR barcodes $secret$. Equation (4) is used to calculate the secret payload.

$$secret = \left\lfloor \frac{ecc}{2} \right\rfloor \times 8 \tag{4}$$

Equation (4) is the traditional method to compute the maximum capacity of the secret payload for QR barcodes. Since the error correction of QR barcodes is individually calculated in groups, the Equation (4) can be detailed as Equation (5):

$$secret = \sum_{n=1}^{N} \left\lfloor \frac{ecc_n}{2} \right\rfloor \tag{5}$$

Equation (5) is the detailed equation of Equation (4). In Equation (5), the value of ecc_n is the number of error-correction codes in group n, and N is the number of groups in the QR barcode. According to Equation (5), the scheme should evenly distribute the secret payload among the various groups to reach the maximum value. Finally, in this step, we obtain *secret*, the configuration of Q, and G, which is a set of data groups in Q.

Step 2. Encrypt the messages M to be hidden. M is a message set composed of m_i, which is divided according to priority. The encryption algorithm is shown in Algorithm 1. Generate a group of keys based on root key k, and use the group of keys to encrypt the messages by groups. Finally, cipher text C and key group K are the outputs.

Algorithm 1 Pseudocode of the encryption algorithm

Input: A string group to be encrypted *texts* and a root key k.
Output: Cipher text C, key group K.

```
K := [];
C := [];
key := k;
for t in texts do
    key := Hash(key);
    c_i := AES(t, key, MODE_ECB);
    C.append(c_i);
    K.append(key);
end for
return C,K;
```

Step 3. Use dynamic Huffman encoding to compress C. Reed–Solomon codes are used to generate error-correction codes for the compressed data. Finally, the scheme separates the data into groups by every eight bits, and T is obtained.

Step 4. Count character c with the highest frequency in T, and use Equation (6) to get S.

$$s_i = t_i \oplus c, (0 \leqslant i \leqslant len(S)) \tag{6}$$

In Equation (6), s_i is one element in S, t_i is one element in T, and $len(S)$ is the size of S. Finally, S is the output.

Step 5. Take an s from S in order. Take g, which is the embedded least payload from G. Embed s in g using the $\oplus$ method.

Step 6. Repeat Step 5 until all g in G reach the error-correction threshold, or all s in S have been traversed.

Step 7. In the final step, G is arranged according to the configuration of Q obtained in Step 1 to generate a QR barcode containing the secret payload.

In real applications, the encrypted key allocation can be arranged according to user needs.

4.7. Extraction Procedure

Common users and unauthorized users can only retrieve QR data from carrier QR barcode Q_c by barcode readers. To extract embedded secret message M' from carrier QR barcode Q_c, Q_c needs to be extracted according to the following steps:

Step 1. Retrieve QR data from carrier QR barcode Q_c and the configuration of Q_c. According to the arrangement, the binary bits of Q_c are grouped to obtain a set G_{Q_c}.

Step 2. Generate QR barcode Q' using QR data and the configuration of Q_c. QR barcode Q' is the original QR barcode without embedded information. According to the arrangement, the binary bits of Q' are grouped to obtain a set $G_{Q'}$.

Step 3. For $g^i_{Q_c} \in G_{Q_c}$, $g^i_{Q'} \in G_{Q'}$, calculate s'_i using Equation (4).

$$s'_i = g^i_{Q_c} \oplus g^i_{Q'}, (0 \leqslant i \leqslant len(G_{Q_c})) \tag{7}$$

In Equation (4), $len(G_{Q_c})$ is the size of G_{Q_c}. Finally, S' is the output, which is the set of s'_i.

Step 4. For $s'_i \in S'$, xor s'_i with c' to obtain T', where c' is the highest-frequency character.

Step 5. All elements in T are stitched into a string in order. Correct this string and decompress it to obtain the cipher text C'.

Step 6. Finally, use the key k' provided by the user to decrypt C' and obtain the plain text hidden in the QR barcode.

As is evident from the extraction procedure, the successful extraction of a secret message requires the successful extraction of the QR data of the carrier QR barcode, the successful restoration of the cipher text, and the provision of the correct key. Successful extraction of QR data requires that the secret payload does not interfere too much with the carrier QR barcode; thus, EasyStego uses the arrangement of QR barcodes to reduce the interference of the carrier QR barcode. In addition, EasyStego uses Reed–Solomon codes to increase the error-correction capability of the cipher texts. Thus, EasyStego can still work in complex physical environments.

5. Evaluation

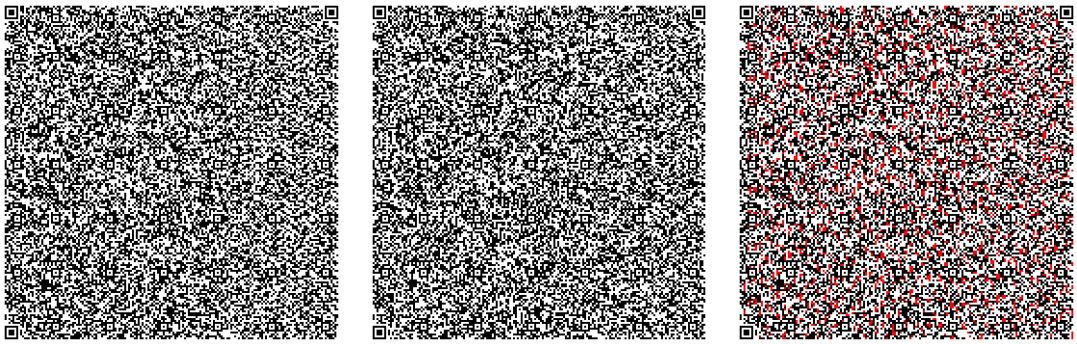
In this section, experiment results and analyses are presented to demonstrate the effectiveness of the proposed scheme. The experiment results answer the following questions:

Q. 1 Can EasyStego cross from the cyberdomain to the physical domain?
Q. 2 What is the maximum capacity of EasyStego?
Q. 3 What is the difference between the secret image of EasyStego and the original image?
Q. 4 How robust is EasyStego?

Section 5.1 answers Q. 1, which proves that EasyStego can cross domains. Section 5.2 answers Q. 2, which proves that EasyStego has high hiding capacity, and the maximum capacity is up to 16,240 bits. Section 5.3 answers Q. 3. It verifies that EasyStego has less impact on the original image than other similar steganography schemes, and secret messages are less likely to be detected. Section 5.4 answers Q. 4 that EasyStego is robust to distortions of complex physical environments.

5.1. Cross-Domain Validity Testing

In this experiment, we demonstrate that EasyStego can cross from the cyberdomain to the physical domain. Additionally, we also tested the cross-domain capabilities of other steganography schemes [4,10,13–15,20,22–24] to illustrate their differences with EasyStego. Figure 2 illustrates a set of samples of the secret embedding procedure. Figure 2a is the original QR barcode. Its QR version is 39, and error-correction level is H. Figure 2b is the Stego QR barcode by the secret embedding procedure. The payload of the embedded secret message is 15,504 bits. Finally, Figure 2c shows the difference between Figure 2a,b.



(**a**) (**b**) (**c**)

Figure 2. Example of an EasyStego-embedded secret message. (**a**) Version 39 ECL:H content: "National University of Defense Technology or People..."; (**b**) QR barcode that embeds a 15,504-bit secret message; (**c**) The difference between (**a**) and (**b**); the change ratio is 15.05%.

The results are shown in Table 4. From left to right in Table 4, the first column lists the steganography scheme, the second indicates the ability to directly extract secret messages in the cyberdomain, and the third indicates the ability to extract secret messages in the physical domain. The results in Table 4 show that EasyStego and References [13,15] can cross domains to extract secret messages from carrier images on printed paper. The reason these works [4,10,20,22–24] cannot cross domains is that their secret payloads are destroyed by physical distortion. The 2LQR [14] uses small textures to embed secret messages. It is difficult for 2LQR to cross domains because the small textures are blurred when the printer prints a secret image onto paper. Another problem with 2LQR is that it is too eye-catching for traditional QR barcodes. This experiment has successfully proven that, in contrast to existing steganography schemes, EasyStego can cross from the cyberdomain to the physical domain.

Table 4. Cross-domain validity testing.

Tool	Cyberdomain	Physical Domain
[4]	✓	×
[24]	✓	×
[10]	✓	×
[20]	✓	×
[22]	✓	×
[23]	✓	×
[14]	✓	×
[15]	✓	✓
[13]	✓	✓
EasyStego	✓	✓

5.2. Steganographic Capacity Testing

In this experiment, we measured the capacity of EasyStego and compared it with References [13,15]. We tested from Version 1 to 40 with different error-correction levels, and 100 iterations were performed for each version. Then, we took the average of the 100 measurements as the capacity of the version.

Measurement results are shown in Figure 3, where the ECL is H. Comparing EasyStego, References [13,15] in Figure 3, we can observe that EasyStego has the largest steganographic capacity of the three, and the difference increases with the version number. Table 5 provides statistical analysis of the EasyStego steganographic capacity after 100 iterations under different QR versions and error-correction levels.

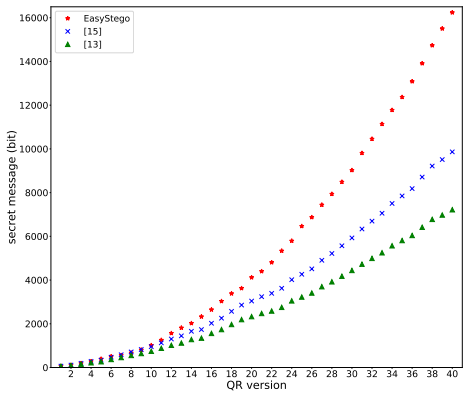


Figure 3. Results of steganographic-capacity testing.

Table 5. Average secret capacity of EasyStego.

Version	Embedded Secret Capacity (bit)			
	L	M	Q	H
1	24	40	48	64
5	104	208	336	392
10	320	608	808	1016
15	568	1088	1904	2328
20	1088	2136	3512	4120
25	1584	3384	5072	6464
30	2496	4864	7472	9088
35	3280	6928	10,278	12,368
40	4432	8736	13,520	16,240

When the version is 40 and the error-correction level is H, the maximum capacity of the EasyStego secret message is 16,240 bits, which is more than the normal QR data-capacity maximum (10,208 bits).

5.3. Difference Testing

To hide information, modifications of carrier QR barcodes are inevitable. This experiment explains the impact of these modifications on carrier QR barcodes and is divided into two parts: the impact on QR data extraction and the difference between carrier QR barcodes and original QR barcodes.

Although EasyStego can hide secret messages without modifying the QR data, it has some effect on scanning and extracting QR barcodes. We measured the impact of EasyStego and References [13,15] on the original QR barcode from the aspects of distance, angle, and error correction. Results are shown in Table 6.

In Table 6, we utilized three different steganography schemes (EasyStego, and References [13,15]) to embed five different sizes of secret messages for the experiment. This table shows the impact of different steganography schemes on QR data-extraction conditions. Here, QR version is 39, and error-correction level is H. In Table 6, the angle is formed by the scanning direction and the normal vector of the plane where the QR barcode is located, distance is the distance from the scan position to the QR barcode, and error pixel is the number of incorrect pixels in the QR barcode. All three schemes have some impact on the extraction of QR data, and the impact of EasyStego is the smallest of the three. Figure 4 shows a comparison of EasyStego and References [13,15] on the change ratio. As shown in Figure 4, the EasyStego change ratio grew the slowest among the three schemes. Thus, when hiding the same size of secret message, EasyStego makes the least modification to data pixels in the original QR barcode.

Table 6. Measurement of the effect on the original QR barcode.

Aspect		5168 bits			9240 bits			9832 bits			15,504 bits		
		[13]	[15]	EasyStego	[13]	[15]	EasyStego	[13]	[15]	EasyStego	[13]	[15]	EasyStego
distance (meter)	0.3	✓	✓	✓	✓	✓	✓	×	✓	✓	×	×	✓
	1	✓	✓	✓	✓	✓	✓	×	✓	✓	×	×	✓
	3	✓	✓	✓	✓	✓	✓	×	✓	✓	×	×	✓
angle (degree)	30	✓	✓	✓	✓	✓	✓	×	✓	✓	×	×	✓
	45	✓	✓	✓	×	✓	✓	×	✓	✓	×	×	✓
	60	✓	✓	✓	×	✓	✓	×	×	✓	×	×	✓
error pixel	10	✓	✓	✓	×	✓	✓	×	×	✓	×	×	✓
	100	✓	✓	✓	×	×	✓	×	×	✓	×	×	×
	400	×	✓	✓	×	×	✓	×	×	✓	×	×	×
change ratio *		8.62%	7.90%	5.66%	15.02%	14.15%	9.53%	N/A **	15.05%	10.07%	N/A **	N/A **	15.01%

- ✓ means that QR data can be extracted, and × means that it cannot be extracted; * Formula for calculating change ratio: $change\ ratio = \frac{ModifiedPixels}{AllDataPixels}$; ** QR barcode is broken and cannot be calculated.

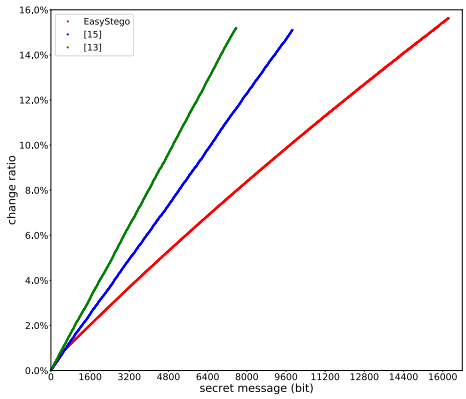


Figure 4. Results of change-ratio testing.

5.4. Robustness Testing

In this experiment, we measured the robustness of EasyStego and compared it with References [13,15]. This experiment measures the robustness of different steganography schemes from the aspects of distances, angles, error correction and texture enhancements. The experiment used standard test image Lena [29] as the texture-enhanced natural image to generate texture-enhanced QR barcodes. As shown in Figure 5, the three QR barcodes had blended textures with different transparencies. As shown in Figure 5, as the texture deepens, the readability of the QR barcode decreases.

(**a**) (**b**) (**c**)

Figure 5. Texture-enhanced QR barcode. (**a**) Texture transparency: 10%; (**b**) Texture transparency: 30%.; (**c**) Texture transparency: 50%.

Results are shown in Table 7, including the cross-domain capabilities of the steganography schemes and robustness in terms of distances, angles, natural textures, and error pixels. Reference [13] can cross domains but cannot resist the impact of angles, textures, and error pixels. Because Reference [13] has no error-correction capability, once any bits of the secret payload are destroyed, it is impossible to extract secret messages from the QR barcode.

Table 7. Robustness comparison of various steganography schemes.

Aspect	[13]	[15]	EasyStego
Cross-domain	✓	✓	✓
Distance	✓	✓	✓
Angle	×	✓	✓
Texture-enhanced	×	Weak	Strong
Error pixel	×	Weak	Strong

Reference [15] can cross domains and has certain error-correction capabilities. However, it is still not easy to extract the secret messages when QR barcodes are imported with natural textures. Reference [15] and EasyStego are used to embed the same secret messages. Then, we used standard test image Lena [29] as the natural texture of images to simulate the effects of light transformation, image wrinkling, and shooting occlusion on the images. After importing natural textures, QR barcodes may incorrectly determine the colors of some pixels during scanning, resulting in an increase in the number of errors. Therefore, it is difficult for Reference [15] to successfully extract secret messages in the real physical domain.

Compared with the above schemes, EasyStego performs well in all measurement aspects. EasyStego can cross domains and resist physical distortions. Additionally, changing the angles and distances of the scan has little impact on EasyStego's extraction. EasyStego is also robust in the impact of natural textures and error pixels. Figure 6 shows the relationship between the number of error pixels and the frequency of successful secret-message extraction. We measured frequency from 0 to 100 bits in intervals of 5. The experiment method was to randomly generate 3000 tests with quantitative error pixels. Measurement objects were QR barcodes embedded with 9240 bits of secret messages by EasyStego, and References [13,15] in version 39 with error-correction level H. From Figure 6, we can see that the frequency of successful extraction by EasyStego, and References [13,15] are 100% at 0 error pixels. The frequency of Reference [15] drops to below 50% when error pixels are increased to 20. When error pixels are increased to 50, the frequency of Reference [15] approaches zero. By contrast, the frequency of successful EasyStego extraction is still near 100%. Even when error bits reach 100 bits, EasyStego always maintains high successful-extraction frequency, approximately 100%.

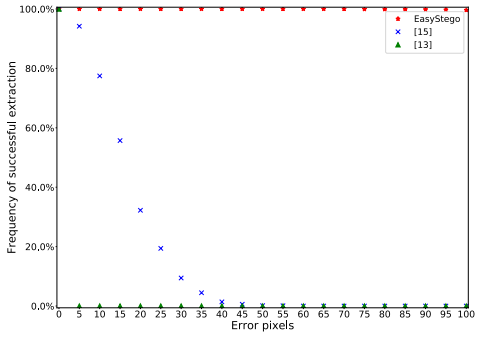


Figure 6. Steganography error correction.

6. Usage Scenarios

In this section, we simulate usage scenarios in the real world. Figure 7 shows a usage scenario of EasyStego when mailing packages. Alice in City *A* wants to send a package to Bob in City *B*. She uses EasyStego to encrypt the information needed to mail the package and hides it in the QR barcode. The generated keys are then uploaded to the Cloud. Key allocation is shown in Table 8. The keys are assigned by the Cloud.

The postman of City *A* decrypts the QR barcode with assigned *key*3 and obtains the destination city of this package.

The postman of City *B* decrypts the QR barcode with assigned *key*2 and obtains the receiver information.

After Bob receives the package, he can decrypt it with assigned *key*1 and obtain the sender information.

Other unrelated people who are not assigned any keys can only obtain the information of the courier company and the alarm call.

This process effectively avoids sensitive-information leakage with regard to senders and recipients. During the transportation process, the bar code appears stained and scratched, but the secret messages can still be extracted successfully.

Table 8. Key allocation.

Name	Value	Data	Holder
root key	Hello,World!	-	Alice
key1	1ef1457579a5b7e8	Name: Alice, Address: City A, Street A, Phone: 01234567.	Bob
key2	eadfef33949752cc	Name: Bob, Phone: 76543210	Postman in City B
key3	bc2cb28527b6f469	Address: City B, Street B.	Postman in City A
no key	-	TEST EXPRESS, LOSS CALL: 400-410073	-

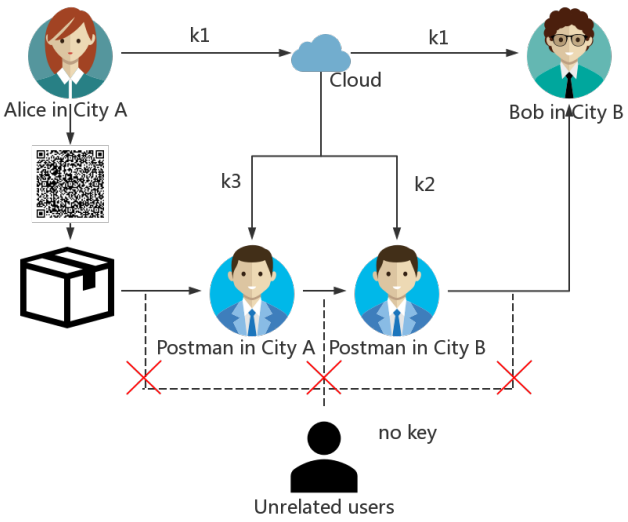


Figure 7. EasyStego usage scenario.

7. Discussion

In this section, we discuss the reasons why EasyStego performs well.

References [13,15], and EasyStego are based on the error-correction capability of QR barcodes in hiding secret messages. In contrast to References [13,15], EasyStego extracts the data arrangement of the QR barcode and embeds the secret payload according to the arrangement. This can reduce the impact of the embedded secret payload on the QR barcode. If the secret payload is blindly embedded without considering the data arrangement of the QR barcode, modifications of eight pixels may result

in the destruction of eight bytes of data. However, according to the data arrangement of QR barcodes to embed the secret payload, the modification of eight pixels needs to only change the data of one byte.

In addition, secret messages are compressed using dynamic Huffman coding. Consequently, a same-sized secret payload can contain more secret messages, which improves the capacity of the steganography scheme.

Last but not least, EasyStego has powerful error-correction capabilities. The error-correction function of EasyStego is mainly derived from two factors: data-part error correction and non-data-part error correction. The location of the embedded secret payload in EasyStego is regular. When data appear in places where they should not appear, the decryption program ignores these errors. In addition, EasyStego uses Reed–Solomon codes to correct the steganographic secret payload, which ensures that the data of the secret payload can still be recovered and decrypted when errors occur.

8. Conclusions

In this paper, we proposed EasyStego, which is a novel cross-domain and antiphysical distortion steganography scheme. EasyStego is based on the use of QR barcodes as carriers; therefore, it is robust to distortions in complex physical environments. EasyStego has large embeddable secret capacity, strong robustness, good concealment, and it can work normally in a variety of situations. We verified the abilities of EasyStego through experiments, and evaluations in real-world scenarios have further demonstrated the effectiveness of our steganography scheme.

Author Contributions: Z.L. contributed to the design of the study, conducted the experiments, analyzed the results, and wrote the manuscript. W.X. analyzed the results and proofread the paper. B.W., Y.T., and Q.X. proofread the paper.

Funding: This research was supported in part by a project of the National Key R and D Program of China under Grant 2017YFB0802300.

Acknowledgments: This study was financially supported by a project of the National Key R and D Program of China under Grant 2017YFB0802300

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AES	Advanced Encryption Standard
QR	Quick Response
DES	Data Encryption Standard
RSA	Rivest–Shamir–Adleman cryptosystem
BMP	Bitmap
DWT	Discrete-Wavelet Transformation
DCT	Discrete-Cosine Transform
EMD	Exploiting Modification Direction
ECC	Error-Correction Codes
ECL	Error-Correction Level

References

1. BBC. Quora Says 100 Million Users Hacked. Available online: https://www.bbc.com/news/technology-46438239 (accessed on 21 December 2018).
2. Guardian, T. Revealed: 50 Million Facebook Profiles Harvested for Cambridge Analytica in Major Data Breach. Available online: https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election (accessed on 21 December 2018).
3. CNET. 150 Million MyFitnessPal Accounts Were Hacked. Here's What to Do. Available online: http://t.cn/EGmveVP (accessed on 21 December 2018).

4. Liao, X.; Yin, J.; Guo, S.; Li, X.; Sangaiah, A.K. Medical JPEG image steganography based on preserving inter-block dependencies. *Comput. Electr. Eng.* **2018**, *67*, 320–329. [CrossRef]
5. Srinivasan, Y.; Nutter, B.; Mitra, S.; Phillips, B.; Ferris, D. Secure transmission of medical records using high capacity steganography. In Proceedings of the 17th IEEE Symposium on Computer-Based Medical Systems, 2004 (CBMS 2004), Bethesda, MD, USA, 24–25 June 2004; pp. 122–127.
6. Castiglione, A.; De Santis, A.; Pizzolante, R.; Castiglione, A.; Loia, V.; Palmieri, F. On the protection of fMRI images in multi-domain environments. In Proceedings of the 2015 IEEE 29th International Conference on Advanced Information Networking and Applications (AINA), Gwangiu, South Korea, 24–27 March 2015; pp. 476–481.
7. Xinhua. China Gets Tough with Personal Information Leaking. Available online: http://www.chinadaily.com.cn/china/2017-05/12/content_29322173.htm (accessed on 24 December 2018).
8. Wave, D. Qrcode. com. QrCode. com [Online]. Available online: http://www.qrcode.com/en/about (accessed on 12 December 2018)).
9. Kumar, R.; Chand, S.; Singh, S. A reversible high capacity data hiding scheme using combinatorial strategy. *Int. J. Multimed. Intell. Secur.* **2018**, *3*, 146–161. [CrossRef]
10. Denemark, T.; Bas, P.; Fridrich, J. Natural Steganography in JPEG Compressed Images. *Electron. Imaging* **2018**, *2018*, 1–10. [CrossRef]
11. Ding, W.; Liu, K.; Yan, X.; Wang, H.; Liu, L.; Gong, Q. An Image Secret Sharing Method Based on Matrix Theory. *Symmetry* **2018**, *10*, 530. [CrossRef]
12. Zhou, X.; Lu, Y.; Yan, X.; Wang, Y.; Liu, L. Lossless and Efficient Polynomial-Based Secret Image Sharing with Reduced Shadow Size. *Symmetry* **2018**, *10*, 249. [CrossRef]
13. Lin, P.Y.; Chen, Y.H.; Lu, E.J.L.; Chen, P.J. Secret hiding mechanism using QR barcode. In Proceedings of the 2013 International Conference on Signal-Image Technology & Internet-Based Systems, Kyoto, Japan, 2–5 December 2013; pp. 22–25.
14. Tkachenko, I.; Puech, W.; Destruel, C.; Strauss, O.; Gaudin, J.M.; Guichard, C. Two-level QR code for private message sharing and document authentication. *IEEE Trans. Inf. Forensics Secur.* **2016**, *11*, 571–583. [CrossRef]
15. Lin, P.Y.; Chen, Y.H. High payload secret hiding technology for QR codes. *EURASIP J. Image Video Process.* **2017**, *2017*, 14. [CrossRef]
16. Petitcolas, F.A.; Anderson, R.J.; Kuhn, M.G. Information hiding-a survey. *Proc. IEEE* **1999**, *87*, 1062–1078. [CrossRef]
17. Mstafa, R.J.; Elleithy, K.M.; Abdelfattah, E. A robust and secure video steganography method in DWT-DCT domains based on multiple object tracking and ecc. *IEEE Access* **2017**, *5*, 5354–5365. [CrossRef]
18. Asad, M.; Gilani, J.; Khalid, A. An enhanced least significant bit modification technique for audio steganography. In Proceedings of the 2011 International Conference on Computer Networks and Information Technology (ICCNIT), Abbottabad, Pakistan, 11–13 July 2011; pp. 143–147.
19. Yang, Z.; Du, X.; Tan, Y.; Huang, Y.; Zhang, Y.J. AAG-Stega: Automatic Audio Generation-based Steganography. *arXiv* **2018**, arXiv:1809.03463.
20. Mehboob, B.; Faruqui, R.A. A stegnography implementation. In Proceedings of the International Symposium on Biometrics and Security Technologies (ISBAST 2008), Islamabad, Pakistan, 23–24 April 2008; pp. 1–5.
21. Lu, T.C.; Tseng, C.Y.; Wu, J.H. Dual imaging-based reversible hiding technique using LSB matching. *Signal Process.* **2015**, *108*, 77–89. [CrossRef]
22. Bansod, S.P.; Mane, V.M.; Ragha, R. Modified BPCS steganography using Hybrid cryptography for improving data embedding capacity. In Proceedings of the 2012 International Conference on Communication, Information & Computing Technology (ICCICT), Mumbai, India, 19–20 October 2012; pp. 1–6.
23. Ramesh, M.; Prabakaran, G.; Bhavani, R. QR-DWT code image steganography. *Int. J. Comput. Intell. Inform.* **2013**, *3*, 9–13.
24. Islam, M.W.; alZahir, S. A novel QR code guided image stenographic technique. In Proceedings of the 2013 IEEE International Conference on Consumer Electronics (ICCE), Las Vegas, NV, USA, 11–14 January 2013; pp. 586–587.
25. BR, R.S.; Rohith, J.; Mukund, V.; Honwade, R.; Rangaswamy, S. Steganography using sudoku puzzle. In Proceedings of the 2009 International Conference on Advances in Recent Technologies in Communication and Computing, Kottayam, Kerala, India, 27–28 October 2009; pp. 623–626.

26. Wave, D. *Information Technology Automatic Identification and Data Capture Techniques QR Code Bar Code Symbology Specification*; International Organization for Standardization, ISO/IEC:18004, February 2015. Available online: https://www.iso.org/standard/62021.html (accessed on 24 December 2018).
27. Huffman, D.A. A method for the construction of minimum-redundancy codes. *Proc. IRE* **1952**, *40*, 1098–1101. [CrossRef]
28. Reed, I.S.; Solomon, G. Polynomial codes over certain finite fields. *J. Soc. Ind. Appl. Math.* **1960**, *8*, 300–304. [CrossRef]
29. Wakin, M. Standard Test Images-Lena. Standard Test Images. Available online: https://www.ece.rice.edu/~wakin/images (accessed on 16 December 2018).