

Article

Spectral Domain-Based Data-Embedding Mechanisms for Display-to-Camera Communication

Lakpa Dorje Tamang and Byung Wook Kim *

Department of Information and Communication Engineering, Changwon National University,
Changwon 51140, Korea; l.tamang9@changwon.ac.kr

* Correspondence: bwkim@changwon.ac.kr

Abstract: Recently, digital displays and cameras have been extensively used as new data transmission and reception devices in conjunction with optical camera communication (OCC) technology. This paper presents three types of frequency-based data-embedding mechanisms for a display-to-camera (D2C) communication system, in which a commercial digital display transmits information and an off-the-shelf smartphone camera receives it. For the spectral embedding, sub-band coefficients obtained from a discrete cosine transform (DCT) image and predetermined embedding factors of three embedding mechanisms are used. This allows the data to be recovered from several types of noises induced in wireless optical channels, such as analog-to-digital (A/D) and digital-to-analog (D/A) conversion, rotation, scaling, and translation (RST) effects, while also maintaining the image quality to normal human eyes. We performed extensive simulations and real-world D2C experiments using several performance metrics. Through the analysis of the experimental results, it was shown that the proposed method can be considered as a suitable candidate for the D2C system in terms of the achievable data rate (ADR), peak signal-to-noise ratio (PSNR), and the bit error rate (BER).

Citation: Tamang, L.D.; Kim, B.W.; Spectral Domain Based Data-Embedding Mechanisms for Display to Camera Communications. *Electronics* **2021**, *10*, 468. <https://doi.org/10.3390/electronics10040468>

Keywords: visible light communication; display-to-camera communication; data embedding; discrete cosine transform

Academic Editors: Jose A. Rabadan, Victor Guerra and Julio Rufo
Received: 19 January 2021
Accepted: 12 February 2021
Published: 15 February 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The recent increase in the use of light-emitting diodes (LEDs) has encouraged research on real-time wireless communication between two devices both equipped with an optical source and image sensor. Optical camera communication (OCC), [1–3] which uses camera image sensors to receive light signals, has acquired important relevance within the visible light communication (VLC) [4–6] field due to its unique feature of ability to separate incident light into spatial and color domains. This trend has attracted researchers and industries for its potential application in overtaking traditional radio frequency (RF) communications, given that the OCC bandwidth requirement and installation cost are lower than those of RF systems.

Display-to-camera (D2C) communication falls under the OCC category where the information is sent from an electronic display to a camera-equipped receiver. Each display pixel serves as a transmitting antenna, and the camera image sensor captures millions of pixels to retrieve the transmitted information. Given the increasing use of multimedia and ubiquitous computing, D2C technology could facilitate the transfer of various types of information in a line-of-sight (LOS) environment, where seamless services are routinely required. In a D2C system, the amount of embedded data is irrespective to the size of the display, and the inserted data should not affect the viewing experience of the user. In addition, data retrieval problems in the receiver must be dealt with, even when the information is sent over a noisy, wireless optical channel. The D2C

system that satisfies these requirements is a promising candidate for 6G communications in addition to RF-based communications for all possible device-to-access networks.

Previous D2C approaches [7–10] mainly considered data embedding in the image spatial domain. In [7], the alpha channel was used to embed data into the translucency change of image pixels; the pixel RGB values were kept unchanged. The INFRAME [8] system enabled “dual-mode full-frame communication” between devices and the human eye. Here, data multiplexing into the full frame of a display unit was achieved by exploiting the concept of complementary frames, where two frames were mutually complementary in terms of the pixel luminance level. Similarly, in the DisCo system [9], data were transmitted by temporally modulating the brightness of the digital display at a very high frequency. Using a rolling shutter camera, transmitted messages were acquired by converting the temporally modulated incident light into a spatial flicker pattern. The SoftLight system [10] used a channel-coding approach that automatically adapts the data transmission rate to different scenarios; suggestions on how to exploit expanded color modulation and coding of bit-level rate-less transmissions of low computational complexity were provided. Quick response (QR) codes [11] and two-dimensional (2D) color barcodes [12, 13] are also considered as a D2C technology where barcoding can be used to embed data. In the COBRA system [12], data were encoded using specially designed 2D color barcodes and streamed between a small screen and the low-speed camera of a smartphone. The size and layout of the code blocks were designed to deal with image blurring during capture in mobile environments; the system allowed for real-time decoding and image processing. Similarly, in the Rainbar system [13], a high-capacity barcode layout allows for flexible frame synchronization and accurate code extraction. Here, to enhance the robustness of the system, “progressive code locator detection” and color recognition were incorporated.

Generally, spatial embedding is vulnerable to various circumstances occurring in a wireless optical channel. The spatial component of an image is described by the pixel intensity, which can be severely affected by ambient lighting and noise present in the wireless optical channel. This noise alters the color intensity of image pixels. The variation in pixel intensity severely compromises the embedded data, which therefore must be explicitly placed in perceptually significant components of the image. When an image is transformed into the spectral domain, its frequency coefficients provide vital information about the image. These coefficients determine the rate of change in pixel values in the spatial domain (i.e., the edges and smooth regions of the image). Therefore, instead of direct manipulation of image pixels, it is preferred to insert data into the coefficients that are robust under all available frequency coefficients. In this manner, the data are spread over many frequency coefficients. The data energy associated with a single coefficient becomes small and thus is not readily detectable in the spatial domain. This renders the data less susceptible to the noise while being transmitted over a wireless optical channel. This situation is analogous to that of spread-spectrum communications, which states that the signal energy at any single frequency is completely undetectable when any narrowband signal is transmitted over a massive bandwidth. Thus, frequency-based embedding is an attractive method to robustly embed data into an image for a D2C link. To embed data, the various types of frequency transformation such as Fourier, wavelet, and cosine transform are considered. Since the maintenance of visual quality of the data embedded image is crucial, the frequency coefficients must be selected such that the visual content of the original image remains unaffected despite data embedding. This topic was previously studied in the context of watermarking and steganography [14–16]. In [14], a robust hybrid watermarking algorithm was presented. Here, a discrete Wavelet transform (DWT) was combined with singular value decomposition (SVD); data were embedded into the DC coefficients of the high-low (HL) and low-high (LH) sub-bands. Similarly, in [15], a fast-discrete Cosine transform (FDCT) algorithm was used to enhance the computational performance of watermarking. Likewise, [16] presented a secure algorithm for image watermarking that

can be generalized to multimedia data, such as audio files and video clips. The embedding principle of digital watermarking is identical to that of D2C frequency-based data embedding. However, conventional watermarking does not consider the effect of the wireless optical channel, which greatly affects the reliability of communication performance.

Previous studies [17–19] explored various methods by which data could be embedded into the frequency domain of an image; a display and a camera served as the transmitter and receiver respectively. In [17], discrete Fourier transform (DFT) coefficients were used to embed data by exploiting wireless orthogonal frequency division multiplexing (OFDM) modulation. The effects of perspective distortion of the transmitted image on the performance of the D2C system were evaluated under various vision transformation parameters. Similarly, [18] and [19] explored data embedding into specific frequency regions of an image using a parallel transmission VLC system. Popular transformation methods, including DWT and discrete cosine transform (DCT) [20], were evaluated. However, these studies have not explored how different types of encoding methods can be achieved in the spectral domain images by utilizing specific embedding factors. The details regarding the selection of frequency coefficients, the effects of the transmission distance D , angle of capture AOC, and wireless optical channel on D2C performance were not investigated either. Specifically, the generalization of these approaches with respect to the changes in the testing environment were not examined.

In this paper, we explore various DCT-based data-embedding schemes in a D2C system. Unlike conventional spectral data-embedding schemes that do not consider the effects of the wireless transmission channel, we derive several data-embedding mechanisms using the DCT coefficients of an image. We consider the nature of the optical wireless channel and embed data into the robust spectral coefficients of the image, which is displayed on an electronic screen and captured by a camera at the receiver side. The decoder at the receiving end performs a geometric correction of the captured image, which is followed by a DCT to extract the embedded data from the frequency coefficients of the captured image. To consider the actual implementation of the proposed scheme, we assessed the effects of the optical wireless channel on the communication performance of the D2C system by conducting various experiments according to multiple parameters. Using software simulations and several real-world experiments, we compare and analyze the performance of our proposed data-embedding schemes in terms of the peak signal-to-noise ratio (PSNR), bit error rate (BER), and achievable data rate (ADR). Our work demonstrates successful data embedding and extraction via a D2C link, which facilitates robust data communication without significantly compromising the visual quality of the display. The main contributions of this research can be summarized as follows:

- Three types of spectral data-embedding mechanisms are presented, and their performance in real-world environments are analyzed.
- The effect of the wireless optical channel on the system performance is assessed under various capture orientations of the receiver camera.
- The performance of the proposed system is assessed according to the data embedding in various positions within the mid-frequency (MF) region of a DCT image.

The remaining of the paper is organized as follows. In Section 2, we provide detailed explanation the D2C architecture, with an emphasis on frequency coefficient selection, encoding, and decoding, where the various types of DCT-based data-embedding mechanisms are discussed. Then, Section 3 presents the performance analysis and evaluation of the proposed model based on both simulations and real-world experiments. Finally, the conclusions are reported in Section 4.

2. Display to Camera Communication System

The basic D2C system includes an electronic display on the transmitter side and a commercial camera on the receiver side. A wireless optical channel acts as a medium to transmit information between the two devices. Figure 1 shows a schematic block diagram of the proposed D2C model. First, the input data vector $\mathbf{b}$ is converted into symbol $\mathbf{S}$ via channel-coding and bits mapping. At the same time, the input image $\mathbf{I}$ is subjected to DCT to obtain its frequency coefficients. Subsequently, a coefficient vector $\mathbf{X_F}$ is selected from the DCT image $\mathbf{I_F}$, where $\mathbf{X_F}$ contains the frequency coefficients preferred for embedding. The data are encoded into the coefficient vector using one of three different mechanisms: M1, additive; M2, multiplicative; or M3, exponential. After data embedding is completed, the embedded coefficient vector $\mathbf{X_E}$ and the other frequency coefficients are concurrently transformed back into the spatial domain via the inverse discrete cosine transform (IDCT) to obtain the data-embedded image $\mathbf{I_E}$. The image quality of $\mathbf{I_E}$ is preserved despite the insertion of additional data; thus, the visual difference between $\mathbf{I_E}$ and $\mathbf{I_{ref}}$ is negligible. Later, $\mathbf{I_E}$ and the reference image $\mathbf{I_{ref}}$ are sequentially transmitted via the electronic screen as shown in Figure 2. These image frame sequences are displayed on the screen as a visual content for the human eye. The human vision system has a temporal resolution of 40–50 Hz, beyond which the flickers cannot be perceived [8]. Taking this into consideration, the transmit rate of the display, $\mathbf{R_D}$ must be higher than the maximum threshold of the temporal resolution of human vision. Therefore, the image frames are sequentially transmitted through the display without any noticeable artifacts or flickers. Note that the purpose of sending $\mathbf{I_{ref}}$, which is adjointly placed with $\mathbf{I_E}$, is to support the decoding at the receiver. The frequency coefficients of the reference image are used to extract the embedded data from the frequency coefficients of the received data-embedded image.

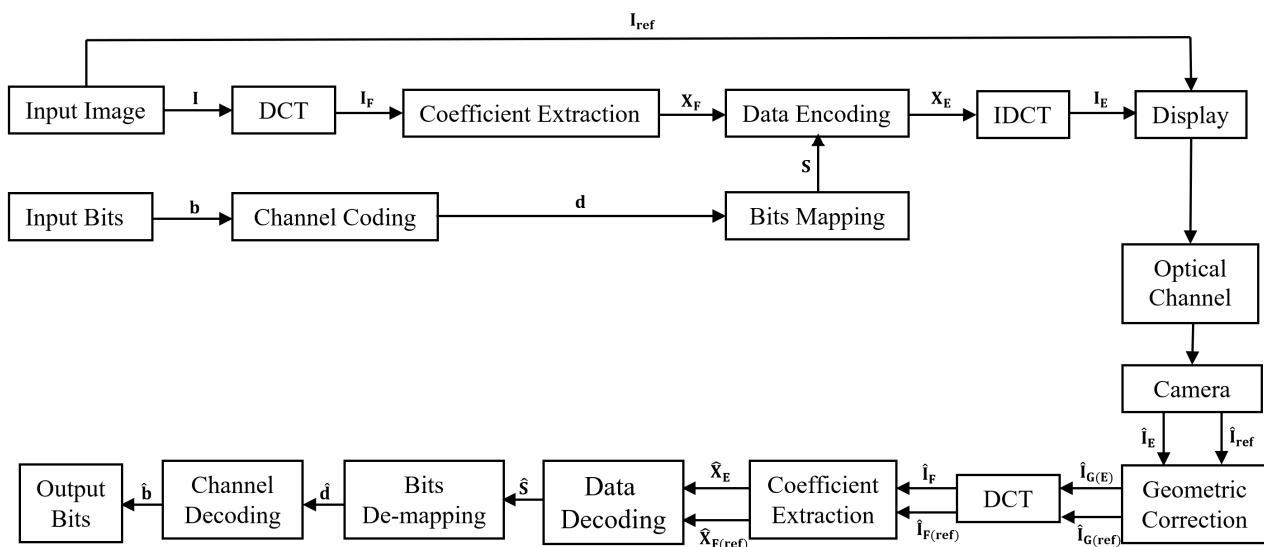


Figure 1. Schematic of the network architecture of the proposed D2C system.

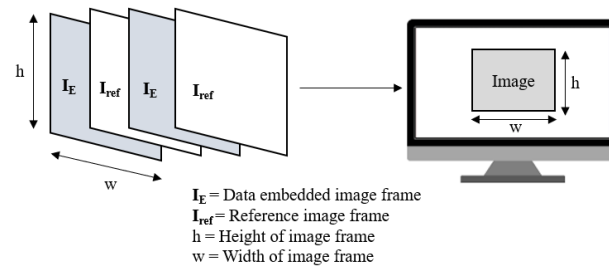


Figure 2. Sequential image frames of the displayed by the screen.

At the receiver side, a rolling shutter camera captures the intensity of display pixels that have passed through a noisy optical wireless channel. Due to the rolling shutter effect, a camera cannot capture all the transmitted image frames if R_D is faster than the capture rate, R_C , of the camera [8]. Therefore, frame synchronization is achieved by capturing the transmitted image frames at least twice; i.e., $R_D \leq \frac{1}{2} R_C$. By doing so, the camera is able to capture all the displayed frames without losing any of them, and the decoding becomes easier. The noises in the channel affect the pixel intensity of the image, which varies markedly from the original ones. The captured images $\hat{I}_E$ and $\hat{I}_{ref}$ are subjected to a degree of geometric distortion and are susceptible to other distortions such as rotation, scaling, and translation (RST) effects. Therefore, a geometric correction technique is utilized to correct the image distortions. As the data are present in the frequency coefficients of the image, the corrected images $\hat{I}_{G(E)}$ and $\hat{I}_{G(ref)}$ are subjected to DCT. Next, the coefficient vectors $\hat{X}_F$ and $\hat{X}_{F(ref)}$ are extracted from the frequency-domain images $\hat{I}_F$ and $\hat{I}_{F(ref)}$, respectively, and used for data decoding. The frequency coefficients of the reference image are simply subtracted from those of the data-embedded image. In this manner, the data can be extracted in the form of symbols $\hat{S}$, which are de-mapped and channel-decoded to obtain the output message in the form of a binary data vector $\hat{b}$.

2.1. Frequency Coefficient Extraction

The entire image requires DCT prior to embedding data into the frequency coefficients of I . Figure 3 shows the positions of the DCT coefficients for an image of size $M \times N$. The top left corner is the DC component representing the average color of the entire frequency-transformed region. The remaining coefficients represent the changes in color when moving from the top left to the bottom right corner of the DCT transformed image. As the DC component preserves the average color of the entire DCT image, data embedding in this region can cause visual artifacts noticeable to the normal human eye. Embedding data in the low-frequency (LF) region (with large-valued coefficients) degrades the visual quality of the image. Likewise, embedding data in the high-frequency (HF) region (with small-valued coefficients) renders the embedded data susceptible to noise present in the channel. This is because JPEG compression in the image capture process is associated with large data loss in the HF spectral region. Therefore, to ensure robust communication in the presence of various types of noises in the optical wireless channel, we embed the data into the MF coefficients. The MF regions are less perceptually significant and carry comparatively less vital information of the image than the HF and LF regions. Therefore, embedding data in this region does not noticeably impact image quality or communication performance.

To select the MF coefficients, all elements from the $M \times N$ -sized I_F matrix are arranged in descending order. Then, after excluding the first 10,000 coefficients underlying in the DC and LF region, the next m highest coefficients are selected for data embedding; m is the total number of input bits. Note that the number of excluded coefficients that occupy the DC component and the LF regions was determined according to the analysis of the value distribution of the frequency coefficients. The

selected m highest coefficients positioned under the MF region are termed the coefficient vector $\mathbf{X}_F$ and modeled as follows:

$$\mathbf{X}_F = [x_1, x_2, x_3, \dots, x_m]. \quad (1)$$

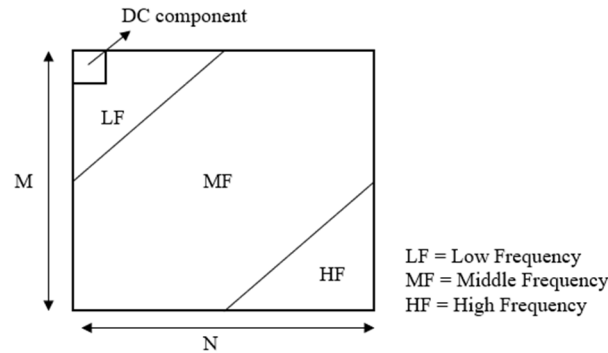


Figure 3. Discrete cosine transform (DCT) coefficients of an image of size $M \times N$.

2.2. Encoding

Before encoding, the input binary data $\mathbf{b} \in \{0,1\}^m$ are first channel-coded and then mapped to the embedding symbols $\mathbf{S}$ of length n . Subsequently, the data bits are mapped to two symbol levels, which are represented as:

$$\mathbf{S} = \begin{cases} -1, & \forall \mathbf{b} : \mathbf{b} = 0 \\ 1, & \forall \mathbf{b} : \mathbf{b} = 1 \end{cases}. \quad (2)$$

Utilizing both $\mathbf{X}_F$ and $\mathbf{S}$ from Equations (1) and (2), we applied three different spectral data-embedding mechanisms: M1, additive; M2, multiplicative; and M3, exponential. They are represented by the equations below:

$$\mathbf{X}_E = \mathbf{X}_F + \alpha \mathbf{S} \quad (3)$$

$$\mathbf{X}_E = \mathbf{X}_F(1 + \alpha \mathbf{S}) \quad (4)$$

$$\mathbf{X}_E = \mathbf{X}_F \times e^{\alpha \mathbf{S}} \quad (5)$$

where $\mathbf{X}_F \in \mathbb{R}^{1 \times m}$ is the MF coefficient vector selected for embedding, as shown in Equation (1). $\mathbf{X}_E \in \mathbb{R}^{1 \times m}$ is the coefficient vector after data embedding, and α is a predetermined embedding factor that affects the visual quality of the data embedded image. More generally, α can be viewed as a relative measure of the amount of adjustment needed to alter the visual quality of the image. We denote the three different mechanisms as follows: M1, additive embedding; M2, multiplicative embedding; and M3, exponential embedding. M1 is the most general type of data-embedding technique and is widely used for digital image watermarking and steganography. In such cases, the coefficient vector is directly added to the product of α and $\mathbf{S}$. M2 is a special case of M1, where $\alpha = \alpha \mathbf{X}_F$, and M3 can be expressed as $\log(\mathbf{X}_E) = \log(\mathbf{X}_F) + \alpha \mathbf{S}$. This is equivalent to M1 but uses the logarithms of the original values of the coefficient vectors.

As the mathematical forms of data-embedding mechanisms differ, their values of α also differ. For M1, a large α is preferred for encoding, whereas M2 and M3 use comparatively smaller values. This is because α variation for M1 minimally impacts the embedding results but significantly affects the results of M2 and M3.

Figure 4 presents the data-embedded images obtained using the three different mechanisms. Figure 4a is the original image, and Figure 4b is the data-embedded image using M1 with a relatively large value of α . Figure 4c,d are data-embedded images using M2 and M3, which are respectively encoded with small values of α . Although their PSNR values differ, encoded images do not contain any obvious artifact when viewed

by the normal human eyes. Thus, the proposed data-embedding method of a D2C system can transmit data while maintaining the original purpose of the display.



Figure 4. Data-embedded images obtained using different mechanisms. (a) Original input image, (b) M1: peak signal-to-noise ratio (PSNR) = 48 dB ($\alpha = 10$), (c) M2: PSNR = 26 dB ($\alpha = 0.1$), and (d) M3: PSNR = 25 dB ($\alpha = 0.1$).

2.3. The Wireless Optical Channel

D2C systems operating in the visible light spectrum (380 to 740 nm) use light-emitting sources to transmit information. These systems operate using a wireless optical channel, where transmission occurs over an unguided medium via visible optical carriers. The wireless optical channel is relatively secure and immune to electromagnetic interference. In addition, the optical channel does not emit harmful radiation, and it operates on inexpensive and readily available transceivers.

The communication performance of a D2C system significantly depends on the orientation and position of the transceiver components, i.e., the AOC and D . The AOC is the angle at which the display plane and camera are mutually aligned. An AOC of 90° affords excellent D2C communication, while a large D between the display and camera limits the image resolution at the receiver. Furthermore, the performance of a D2C system is seriously affected by interference from opaque objects in the path of the visual components.

In a D2C system, the information signal is modulated by the light intensity. This is termed intensity-modulated direct detection (IM/DD), where the transmitted signal is proportional to the light intensity. Similarly, on the receiver side, the camera captures the pixel intensity of the display. The pixel is the primary unit for detection of received power, which corresponds to the photon arrival rate at the receiver. In OCC, the photon arrival rate is usually considered to be stochastic and is modeled using Poisson's distribution. Therefore, the signal received via an optical wireless channel is affected by several types of noise. The amount of noise captured by the receiver, along with the display content, is affected principally by the room configuration, reflective characteristics of the display, AOC, transmission range, and position and orientation of the receiver. For an indoor D2C system, performance is limited by background illumination and external light sources, such as bulbs and incandescent lamps. Such sources induce a constant shot noise element at the receiver side, thus degrading the overall communication performance.

Although both the display and camera are digital devices, the transmission medium is analog and completely wireless. This means that digital-to-analog (D/A) and analog-to-digital (A/D) conversion are performed during both transmission and reception,

which tends to degrade the transmitted signals. In addition, depending on the capture position of the camera, complex phenomena such as blurring and geometrical distortion may be evident in a captured image. Since a D2C system should consider all these operating conditions of the optical wireless channel, it differs from traditional image watermarking systems.

2.4. Decoding

The receiver camera captures the transmitted image sequence while looking at the electronic display. The captured images $\hat{\mathbf{I}}_{\mathbf{E}}$ and $\hat{\mathbf{I}}_{\text{ref}}$ undergo geometric correction to nullify spatial domain distortions, such as geometrical distortions and RST effects instigated by the transmission channel. Figure 5a,b show the captured image $\hat{\mathbf{I}}_{\mathbf{E}}$ and reconstructed image $\hat{\mathbf{I}}_{\mathbf{G}(\mathbf{E})}$ after geometric correction, respectively. The corners of $\hat{\mathbf{I}}_{\mathbf{E}}$ are irregular, and the image is geometrically distorted. For geometric correction, we employed a perspective transformation technique using the four corners of the region of interest (ROI) in the captured image. This process was also applied to the captured reference image $\hat{\mathbf{I}}_{\text{ref}}$ to generate the same-sized image. Next, the corrected images $\hat{\mathbf{I}}_{\mathbf{G}(\mathbf{E})}$ and $\hat{\mathbf{I}}_{\mathbf{G}(\text{ref})}$ were subjected to DCT and transformed into their respective spectral domains $\hat{\mathbf{I}}_{\mathbf{F}}$ and $\hat{\mathbf{I}}_{\mathbf{F}(\text{ref})}$. From the available coefficients, the MF region was selected in a manner as described in Section 2.1, and the respective coefficient vectors $\hat{\mathbf{X}}_{\mathbf{E}} = [\hat{x}_1, \hat{x}_2, \dots, \hat{x}_m]$ and $\hat{\mathbf{X}}_{\mathbf{F}(\text{ref})} = [\hat{x}_{\mathbf{F}(\text{ref})}^1, \hat{x}_{\mathbf{F}(\text{ref})}^2, \dots, \hat{x}_{\mathbf{F}(\text{ref})}^m]$ were extracted. Finally, by subtracting $\hat{\mathbf{X}}_{\mathbf{F}(\text{ref})}$ from $\hat{\mathbf{X}}_{\mathbf{E}}$, the data were extracted in the form of symbols as:

$$\hat{\mathbf{S}} = \hat{\mathbf{X}}_{\mathbf{E}} - \hat{\mathbf{X}}_{\mathbf{F}(\text{ref})}. \quad (6)$$

The symbol $\hat{\mathbf{S}}$ obtained from the above simple subtraction contains information on the optical wireless channel of the D2C link, the original signal, and the embedding strength. Since the original symbol $\mathbf{S}$ has levels of -1 and 1 , the decoding process can be performed through the sign of $\hat{\mathbf{S}}$. The acquired symbol is a real number and is de-mapped to binary values $\hat{\mathbf{d}}$ and channel-decoded into output message vector $\hat{\mathbf{b}} \in \{0, 1\}^m$, as follows:

$$\hat{\mathbf{b}} = \begin{cases} 1, & \forall \hat{\mathbf{S}} : \hat{\mathbf{S}} \leq 0 \\ 0, & \forall \hat{\mathbf{S}} : \hat{\mathbf{S}} > 0 \end{cases}. \quad (7)$$

Note that the proposed decoding method requires a very simple operation; thus, it can be easily applied to a communication service requiring a small-scale and complexity-constrained receiver hardware.

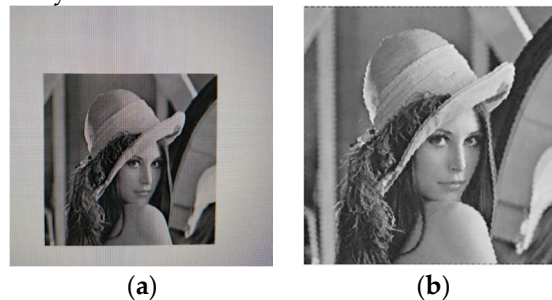


Figure 5. (a) The captured image, and (b) reconstructed image

3. Performance evaluation and results

In this section, we describe the results of simulations and real-world experiments of the proposed model, and performance evaluations in terms of BER, PSNR, and ADR. The real-world experiment was performed in an indoor environment considering a wireless optical channel between the display and camera. The parameters used in the real-world experiments are shown in Table 1. The frame synchronization problem was alleviated by setting the capture rate of the camera to twice the transmitting rate of the display. To observe the feasibility of the proposed model with varying input data size,

we set the length of data vector to 250 and 500 bits. Embedding was performed in such a way that all bits were successfully encoded in the preferred position of spectral domain image, signifying a 100% embedding rate. Figure 6 shows actual experimental environments under various lighting conditions. Figure 6a–c show experiments performed under normal lighting conditions with $D = 15$ cm, $AOC = 90^\circ$; $D = 20$ cm, $AOC = 90^\circ$; and $D = 15$ cm, $AOC = 30^\circ$, respectively. Similarly, Figure 6d shows an experimental environment in the presence of ambient lighting with D and AOC set to 20 cm and 90° , respectively. For the input image, a 256×256 grayscale Lenna image was used to embed the data. The data bits were encoded using the three different embedding mechanisms described in Section 2.2. For M1, α was set in the range $[10, 50]$, and for M2 and M3, the α values were in the range $[0.1, 0.3]$. For error correction, we used the $\frac{1}{2}$ convolution code in the experiment.

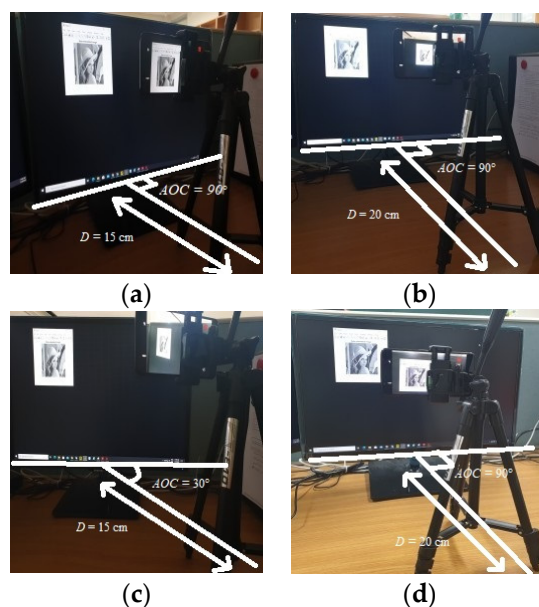


Figure 6. Experiments conducted under various lighting conditions. (a) Normal lighting, (b) normal lighting with a varying distance, (c) normal lighting with a varying angle, and (d) ambient lighting.

Table 1. Experimental parameters

Parameters	Values
Binary data bits	250, 500
Image	Lenna grayscale (256×256)
Display unit	Samsung monitor, 59.8 cm (1920×1080), transmit rate of 60 Hz
Receiver camera	One-plus 5T, 20 Megapixel, (1920×1080), capture rate of 120 fps
Test environment	Indoor

3.1. BER vs PSNR

Figure 7a presents the simulated BER performance of all three data-embedding mechanisms with respect to the PSNR values. For this simulation, we considered the scenario in which the data-embedded images were not affected by channel noise. To observe the effect of geometric image rotation on BER performance, we rotated the data-embedded images by 10° . In all three mechanisms, we can observe the degradation of BER performance with increasing PSNR. However, at similar BER values, M1 outperformed M2 and M3 in terms of PSNR. When data were directly added to the

selected frequency coefficients, the intensity of image pixel was less affected, and thus the PSNR of M1 was better than others. Furthermore, we can observe a significant increment of BER for all mechanisms when the PSNR increased along the x-axis. Although a higher PSNR indicates better image quality, it is achieved in the cost of degraded communication performance. From Equations (3), (4), and (5), we can note that the PSNR has an inverse relationship with α . With a small value of α , the visual quality of the data-embedded image is well preserved, and the PSNR value is high. On the other hand, a larger α tends to degrade image quality and the PSNR value becomes low. In Figure 7b, when the size of $\mathbf{b}$ was increased from 250 to 500, we can see that BER performance is slightly degraded. However, since this result is caused by sending data twice as much, the user can control this trade-off relationship according to the communication requirements of the target system.

Figure 8 depicts the BER performance under real-world conditions with optical wireless channel, using a commercial display and off-the-shelf smartphone camera. It can be seen from the figure that the trend of the overall curves is similar to that of Figure 7. For M1, M2, and M3, the BER is observed to be poorer for a higher PSNR level, resulting in a degradation of communication performance. As a wireless optical channel was introduced between the display and camera during the real-world experiment, the system was more resilient to errors, leading to a much greater deterioration of the BER than in the previous experimental results of Figure 7.

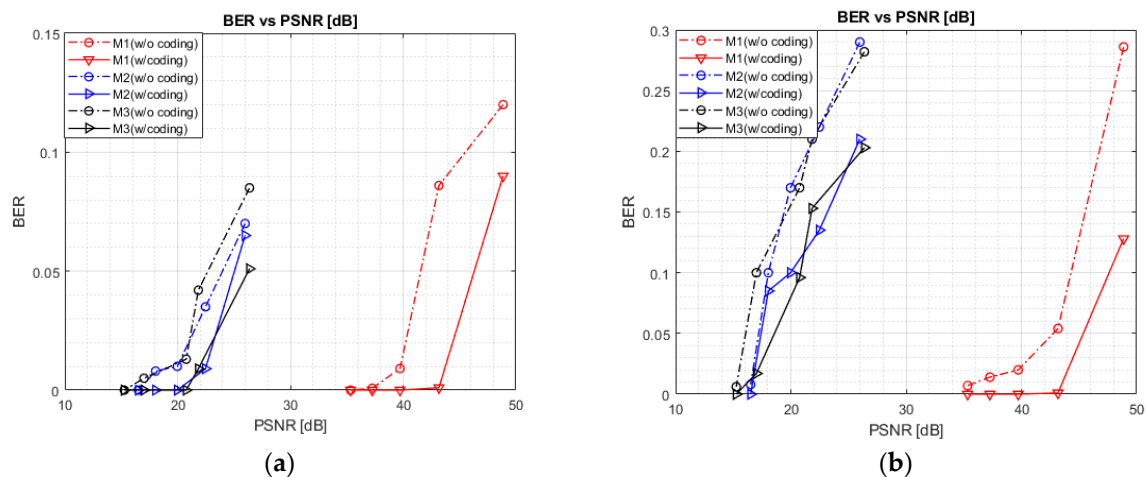


Figure 7. Bit error rate (BER) performance with respect to PSNR [dB], in simulations in which images were rotated at the receiver. (a) $m = 250$ and (b) $m = 500$.

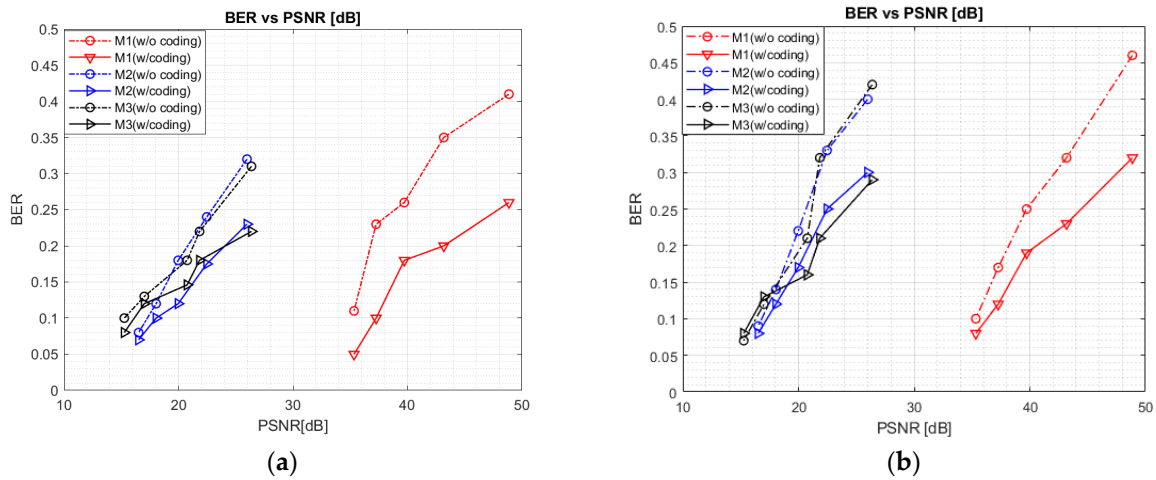


Figure 8. BER performance with respect to PSNR [dB] for images captured via an optical wireless channel. (a) $m = 250$ and (b) $m = 500$.

3.2. ADR vs PSNR

Figure 9 presents the ADR of the proposed D2C model with respect to the PSNR values for all three mechanisms of data embedding. We performed a real-world experiment and image rotation simulation respectively, with and without considering the optical wireless channel. We can observe that the system yields relatively higher ADR values at a lower PSNR level. This is because the ADR is directly related to the BER, i.e., $(1 - BER) \times m$. For all three data-embedding mechanisms, a maximum ADR of approximately 460 bps was achieved under real-world conditions for all perspective parameters.

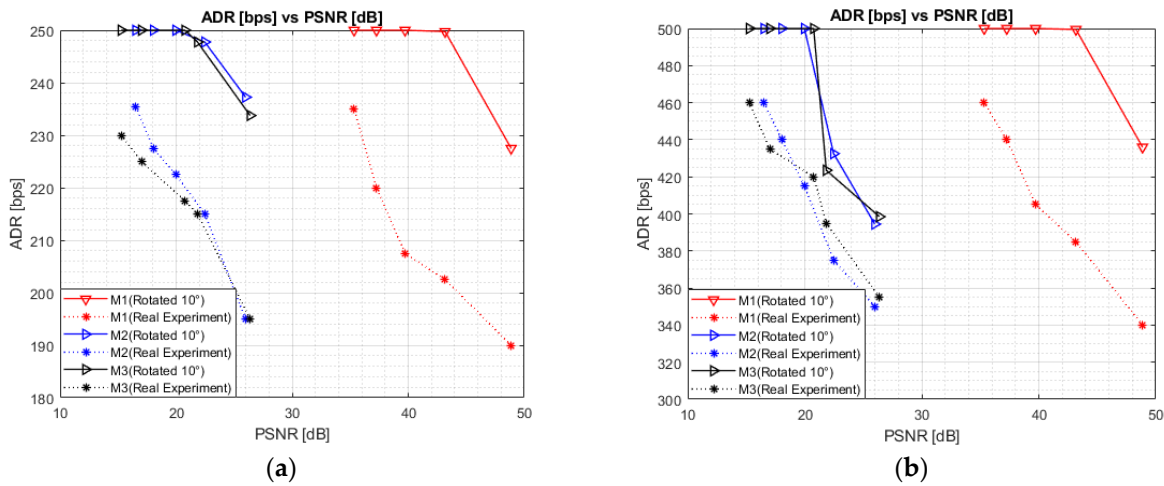


Figure 9. Achievable data rate (ADR) [bps] against the PSNR [dB] performance of the proposed system. (a) $m = 250$ and (b) $m = 500$.

3.3. Transmission Distance and the AOC

Figure 10 shows the BER performance of all three data-embedding mechanisms with respect to D and AOC. For the experiment, α for M1 was set to 30, and that for M2 and M3 was set to 0.2. To investigate the BER performance with distance, D was increased from 10 to 30 cm. The AOC was set to 90° , representing a perfect angular alignment between the display and camera. Figure 10a shows that the BER performance is degraded with increasing D . This is attributable to image blurring and an out-of-focus effect introduced when capturing images over a large distance. Blurring introduces phase noise, which impairs the transmitted data; blur is usually modeled as additive white Gaussian noise (AWGN) in wireless communication scenarios. Similarly, Figure

10b shows the effect of the AOC on the BER performance of the D2C system. To observe the effect of the AOC on communication performance, we varied the angle from 30° to 150° at a constant transmission distance of 15 cm. We can observe an excellent BER performance for an AOC of 90° , but errors occurred as the AOC was significantly varied around 90° . BER performance is completely dependent on the transmission distance and angle at which the camera captures the display content. Therefore, to optimize D2C communication performance, the display and camera must be close, and the AOC should be set as close to 90° as possible.

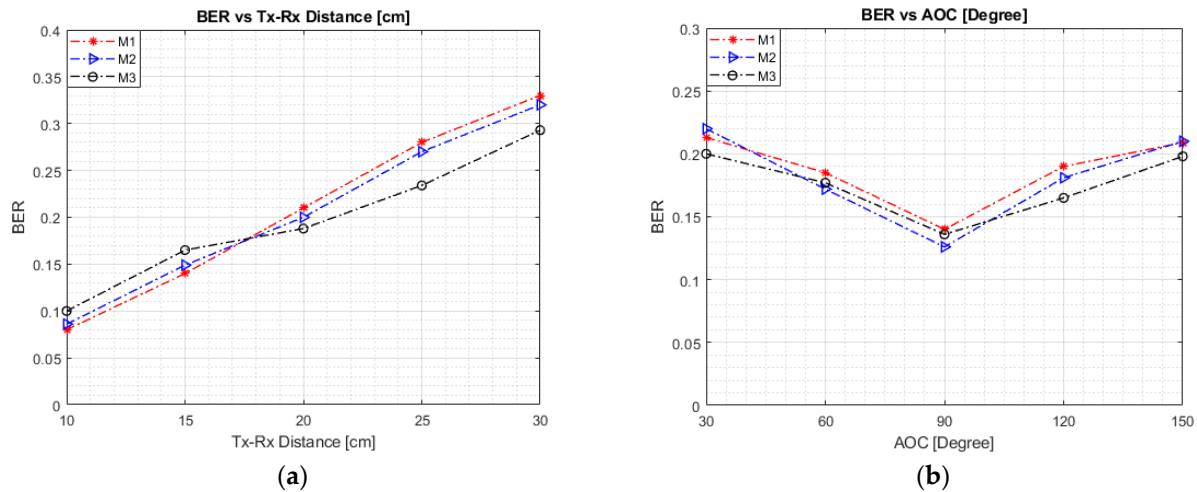


Figure 10. The BER performance of the proposed system (a) by distance (D) and (b) angle of capture (AOC).

3.4. D2C Performance According to Changes in Ambient Lighting

Figure 11 shows the BER performance of the proposed D2C system under ambient lighting. Experiments were performed inside a closed room with “normal” and “ambient” lighting. To achieve normal lighting, the room lights were turned off and only the natural light was provided. On the other hand, under the ambient lighting condition, all room lights were turned on. For all data-embedding mechanisms, under the similar PSNR levels, we can observe that the BER performance is relatively poorer when images are captured with “ambient” lighting compared to “normal” lighting. This is because the captured images under “ambient” lighting were overexposed to light and their intensities were thus much greater than usual. This lowers the probability of successful decoding when data is extracted from the captured image. It is also clear from the figure that of the three data-embedding mechanisms, M1 can provide an optimal BER performance of 0.06 with a PSNR of approximately 35 dB when tested under “ambient” lighting conditions. Therefore, regardless of the transmission environment under unfavorable lighting conditions, it is certainly possible to choose data embedding with M1 in order to achieve robust communication performance and higher image reconstruction quality. Note that the ultimate BER performance of the proposed system under frequently changing lighting conditions is significantly dependent on the amount of variation or alteration in the pixel intensities of the captured image.

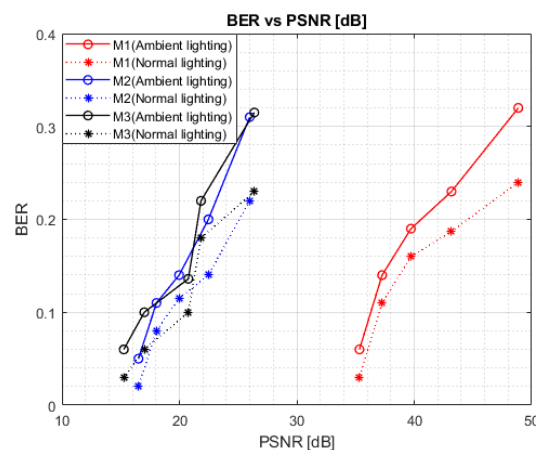


Figure 11. The BER against the PSNR [dB] for the proposed system under normal and ambient lighting conditions.

3.5. Embedding in Different Areas of the MF Region

In this sub-section, we evaluate the performance of the proposed D2C model by embedding data into different areas within the MF region of a DCT image. The MF region was split into three sub-regions, each of which holds 200 bits of $\mathbf{b}$. Here, the α value for M1 was set to 30, and those for M2 and M3 were set to 0.15. D and AOC were set to 15 cm and 90° , respectively. To assign embedding positions for the sub-regions (A, B, and C), the coefficients of the MF region were sorted in descending order. The bit capacity for each sub-region was specified based on the sorted coefficients, where a total of 600 highest coefficients were selected (200 for each sub-region). Sub-region A contained the top 200 coefficients and so on for sub-regions B and C.

Embedding data in different sub-regions did not alter the BER performance of any mechanisms. However, the visual quality of the data-embedded image was affected instead. Table 2 shows the PSNR data of the three data-embedding mechanisms. In each case, three different sub-regions within the MF region were selected for embedding. Here, each sub-region acts as a separate coefficient vector of length 200. We can observe that the PSNR levels of the three sub-regions differed significantly, even though the data-embedding mechanism was identical. M1 exhibited the highest PSNR level when the data were embedded into sub-region C. With M1, the smaller coefficients of this region are directly added to the product of α and $\mathbf{S}$. Whenever such values are selected for embedding, the difference between the input image and data-embedded image is minimal, leading to a better PSNR. On the other hand, for M2, the PSNR increased significantly when data were embedded into sub-region C of the MF region. Here, the smaller coefficients of the sub-region are multiplied by $\alpha\mathbf{S}$, yielding very small differences between the input and data-embedded images, and thus a higher PSNR. Similarly, for M3, the PSNR was highest when sub-region C was selected for embedding. Here, $\alpha\mathbf{S}$ is exponentially multiplied by the selected coefficient vector, which presents also small values. Thus, a higher PSNR can be maintained, given only the slight difference between the input and data-embedded images. Note that these sub-regions tend to reside on the lower part of the MF coefficients, close to the HF region. The insertion of data into such regions prevails successful data embedding with less visual artifacts and the higher PSNR level.

Table 2. PSNR performance of the proposed display-to-camera (D2C) model featuring data embedding in different areas of the mid-frequency (MF) region.

Mechanisms	MF sub-regions	PSNR [dB]
M1: Additive	Sub-region A: 1–200	27.3
	Sub-region B: 201–400	46.8
	Sub-region C: 401–600	53.9
M2: Multiplicative	Sub-region A: 1–200	24.4
	Sub-region B: 201–400	27.3
	Sub-region C: 401–600	28.9
M3: Exponential	Sub-region A: 1–200	26.9
	Sub-region B: 201–400	27.4
	Sub-region C: 401–600	29.5

4. Conclusion

This paper presented three different frequency-based data-embedding mechanisms for a D2C system. By selecting the robust frequency coefficients within the DCT image, data were embedded by using predetermined embedding factors that varied according to the embedding mechanism. To preserve the original purpose of display, all embedding mechanisms were carefully designed so that none of them compromised the viewing experience of the user. The effects of the signal processing properties of transmission channel on D2C system performance were also analyzed through extensive real-world experiments, with several performance metrics that include transmission distances and angles, and data transmission in an ambient lighting environment. The experimental results presented the maximum ADR of approximately 460 bps and a robust PSNR and BER performance for all data-embedding mechanisms within the MF regions of the spectral domain image. Therefore, the proposed scheme can be a viable solution to the D2C system requiring good communication performance in a display-to-camera link while preserving the visual quality of the original image.

Author Contributions: Conceptualization, L.D.T., and B.W.K.; methodology, L.D.T., and B.W.K.; software, L.D.T.; validation, L.D.T.; writing—original draft preparation, L.D.T., and B.W.K.; writing—review and editing, L.D.T., and B.W.K.; supervision, B.W.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by National Research Foundation of Korea (NRF) grant funded by the Korean government (NRF-2019R1A2C4069822)

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

References

1. Saeed, N.; Guo, S.; Park, K.; Al-Naffouri, T.Y.; Alouni, M.S. Optical camera communications: Survey, use cases, challenges, and future trends. *Phys. Commun.* **2019**, *37*, 1–17.
2. Le, N.T.; Jang, Y.M. MIMO architecture for optical camera communications. *J. Korean Inst. Commun. Inf. Sci.* **2017**, *42*, 8–13.
3. Chowdhury, M.J.; Hossan, M.T.; Islam, A.; Jang, Y.M. A comparative survey of optical wireless technologies: Architectures and applications. *IEEE Access* **2018**, *6*, 9819–9840.
4. Rehman, S.; Ullah, S.; Chong, P.H.J.; Yongchareon, S. Visible light communication: A system perspective—Overview and challenges. *Sensors* **2019**, *19*, 1153.
5. Chow, C.W.; Chen, C.Y.; Chen, S.H. Visible light communication using mobile-phone camera with data rate higher than frame rate. *Opt. Express* **2015**, *23*, 26080–26085.
6. Cevik, T.; Yilmaz, S. An overview of visible light communication systems. *Int. J. Comput. Netw. Commun.* **2015**, *7*, 139–150.
7. Li, T.; An, C.; Campbell, A.T.; Zhou, X. HiLight: Hiding bits in pixel translucency changes. In Proceedings of the 1st ACM Workshop on Visible Light Communication Systems, Maui, HI, USA, 7 September 2014; pp. 45–50.
8. Wang, A.; Peng, C.; Zhang, O.; Shen, G.; Zeng, B. InFrame: Multiflexing full-frame visible communication channel for humans and devices. In Proceedings of the HotNets-XIII: Proceedings of the 13th ACM Workshop on Hot Topics in Networks, Los Angeles, CA, USA, 27–28 October 2014; pp. 1–7.

9. Jo, K.; Gupta, M.; Nayar, S.K. DisCo: Display-camera communication using rolling shutter sensors. *ACM Trans. Graph.* **2016**, *35*, 1–13.
10. Du, W.; Liando, J.C.; and Li, M. SoftLight: Adaptive visible light communication over screen-camera links. In Proceedings of the IEEE INFOCOM 2016—The 35th Annual IEEE International Conference on Computer Communications, San Francisco, CA, USA, 10–14 April 2016; pp. 1–9.
11. Lin, S.; Hu, M.; Lee, C.; Lee, T. Efficient QR code beautification with high quality visual content. *IEEE Trans. Multimed.* **2015**, *17*, 1515–1524.
12. Hao, T.; Zhou, R.; Xing, G. COBRA: Color barcode streaming for smartphone systems. In Proceedings of the 10th International Conference on Mobile Systems, Applications, and Services (MobiSys '12). Association for Computing Machinery, New York, NY, USA, 25–29 June 2012; pp. 85–98.
13. Wang, Q.; Zhou, M.; Ren, K.; Lei, T.; Li, J.; Wang, Z. Rain Bar: Robust application-driven visual communication using color barcodes. In Proceedings of the 2015 IEEE 35th International Conference on Distributed Computing Systems, Columbus, OH, USA, 29 June–2 July 2015; pp. 537–546.
14. Zhou, X.; Zhang, H.; Wang, C. A robust image watermarking technique based on DWT, APDCBT, and SVD. *Symmetry* **2018**, *10*, 77.
15. Tsai, S.; Yang, S. A fast DCT algorithm for watermarking in digital signal processor. *Math. Probl. Eng.* **2017**, 1–7, doi:10.1155/2017/7401845.
16. Cox, I.J.; Kilian, J.; Leighton, F.T.; Shamoon, T. Secure spread spectrum watermarking for multimedia. *IEEE Trans. Image Process.* **1997**, *6*, 1673–1687.
17. Kim, B.W.; Kim, H.; Jung, S. Display field communication: Fundamental design and performance analysis. *J. Lightwave Technol.* **2015**, *33*, 5269–5277.
18. Mushu, R.; Wada, T.; Mukumoto, T.; Okada, H. A proposal of information embedding scheme based on discrete cosine transform in parallel transmission visible light communications. In Proceedings of the 2018 IEEE 7th Global Conference on Consumer Electronics (GCCE), Nara, Japan, 9–12 October 2018; pp. 175–176.
19. Mushu, R.; Wada, T.; Mukumoto, T.; Okada, H. A study on information embedding in color images using discrete cosine transform for visible light communications. In Proceedings of the 1st Workshop on Optical Wireless Communication for Smart City, Nagoya, Japan, 17–18 December 2019.
20. Tamang, L.D.; Kim, B.W. Exponential Data Embedding Scheme for Display to Camera Communications. In Proceedings of the 2020 International Conference on Information and Communication Technology Convergence (ICTC), Jeju Island, Korea, 21–23 October 2020; pp. 1570–1573.