MoiréComm: Secure Screen-camera Communication based on Moiré Cryptography

Hao Pan, *Member, IEEE,* Yongjian Fu, *Student Member, IEEE,* Yu Lu, *Student Member, IEEE,*
Feitong Tan, *Member, IEEE,* Yi-Chao Chen, *Member, IEEE,* Ju Ren, *Senior Member, IEEE*

***Abstract*—Quick Response (QR) codes have become increasingly popular for screen-camera communication due to their swift readability and widespread smartphone use. Nevertheless, they are vulnerable to privacy invasions from unauthorized photography. Addressing this, we propose a novel Moiré encryption technique-based secure screen-camera communication system, named MoiréComm. The Moiré encryption can enhance security by using distinct spatial frequency patterns for camouflage. The original QR code is revealed as a Moiré pattern only when the camera in a designated position, *e.g.*, directly in front and 30 cm from the screen. From any other positions, only the camouflaged QR code can be seen. Decryption schemes are customized for different scenarios. The multi-frame approach achieves a decryption success of over 98.6% within 13.2 frames in handheld scenarios. Conditional generative adversarial network (cGAN)-based decryption method decodes the Moiré QR code images with a 98.8% success rate in 0.02 s within three frames and is also applicable in handheld scenarios. For fixed screen-camera setups, our fast decryption scheme achieves 99.4% success within two frames, with average 0.4 s latency. Significantly, the decryption rate plunges to 0% for surveillance cameras displaced by 20° or more than ≥10 cm from the target position, demonstrating MoiréComm's resilience against attacks.**

***Index Terms*—Screen-camera communication, secure QR code, nonlinearity.**

I. Introduction

Screen-camera communication leverages the visible light channel to facilitate data transmission. Quick response (QR) codes are a widely used form of this technology, allowing rapid access to various quickly access specific and user authentication services, such as mobile payments, building access controls, and library book rentals via omnipresent screens and mobile device cameras. Fig. 1 features the Alipay application [1], which serves as an illustrative case for a QR code-based authentication system. In this system, the user initiates a transaction by entering a password, which triggers the application (app) to encrypt the pertinent account and payment data, subsequently generating a QR code on the user's smartphone screen. The code is scanned by a retailer using either an Electronic Cash Register (ECR) or a smartphone, allowing for decryption of the encoded data. A payment request is then dispatched to the Alipay system's backend infrastructure. Once the transaction is validated, the user is informed in accordance with the established transaction procedure [2].

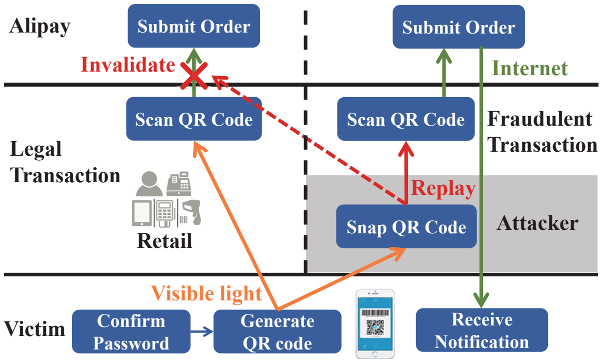
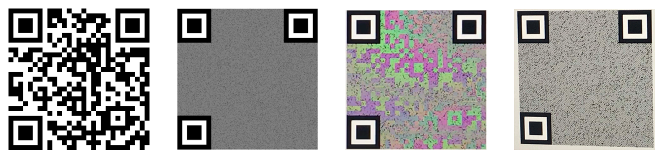


Fig. 1. Security risk of the existing screen-camera communication (*e.g.*, QR code). Left is the QR code-based authentication system for Alipay, and right is the process of STLS attack on QR code-based payment

As a widely used form of screen-camera communication, QR codes offer convenience but are also prone to security vulnerabilities, including Replay attacks [3], [4], [5] and Synchronized Token Lifting and Spending (STLS) attacks [6]. In these scenarios, an adversary can covertly capture a user's QR code and use it for unauthorized payments or to gain access to private information. Also as depicted in Fig. 1, an attacker can simply photograph a victim's QR code and present it to a retailer to conduct an illicit transaction without needing to decrypt the message or acquire the victim's password. The prevalence of QR codes in retail chains (*e.g.*, Walmart and Starbucks), financial institutions (*e.g.*, PayPal and Alipay), and social networking apps (*e.g.*, WeChat and Instagram) exacerbates the potential for substantial financial loss [7], [8] and mass privacy breaches [9], [10]. The complete threat model is outlined in Sec. II.

The existing vulnerabilities in screen-camera systems are attributable to the inherent characteristics of the visible light channels. The light emitted from the screen exhibits hemispherical radiation patterns, allowing the transmission of information to be received at almost any unobstructed location. Moreover, the screen-camera communication systems are deficient in a secure mutual challenge-response protocol due to their unidirectional information flow architecture, which

Corresponding author: Yongjian Fu
H. Pan, is with the Microsoft Research Asia, Shanghai, China.
E-mail: panhao@microsoft.com
Y. Fu is with the Department of Computer Science and Technology, Tsinghua University, Beijing, China.
E-mail: yongjianwork@gmail.com
Y. Lu and Y. Chen are with the College of Computer Science and Technology, Shanghai Jiao Tong University, Shanghai, China.
E-mail: yulu01, yichao@sjtu.edu.cn
F. Tan is with the School of Computing Science, Simon Fraser University, Burnaby, British Columbia, Canada.
E-mail: feitongt@sfu.ca
J. Ren is with the Department of Computer Science and Technology, Tsinghua University, Beijing, China.
E-mail: renju@tsinghua.edu.cn



(a) Original QR code image (b) Encrypted QR code image (c) Picture taken at designated position (d) Picture taken at wrong position (off by 15°)

Fig. 2. Natural revelation of Moiré code upon camera alignment with designated position to the screen and indiscernible grayscale pattern when misaligned

is limited to screen-to-camera communication without an acknowledgment (ACK) feedback loop from the camera to the screen. In this paper, we develop a novel and secure encryption technique for the screen-camera communication systems, hereafter referred to as MoiréComm to resist such attacks.

Our proposed scheme leverages the non-linear spatial frequency characteristics of light to effectively prevent unauthorized access to QR codes within a communication channel. MoiréComm harnesses the existing physical characteristics of both the screen and the camera to establish an encryption protocol, thereby eliminating the need for additional communication channels or specialized hardware components. When a QR code is initially created (refer to Fig. 2(a) for an example), MoiréComm secures it by integrating it into a pattern that the human eye perceives as noise, leveraging a predetermined spatial frequency for this purpose. Fig. 2(b) presents an instance of such an encrypted QR code, hereinafter referred to as Moiré code. Upon display and subsequent capture of this Moiré code by a camera, the resulting image is impressed upon the camera's CMOS sensor, subject to geometric alterations such as scaling, translation, and rotation, which are influenced by the spatial relationship between the screen and the camera. Precise alignment of the camera, acting as the designated receiver, with a specific location allows the nonlinearities of spatial frequency between the projected Moiré code and the camera's color filter array (CFA) to decode the original QR code via a Moiré pattern, as evidenced in Fig. 2(c)[1]. Conversely, any unauthorized cameras attempt to capture the Moiré code from alternative positions, they will fail to discern the correct Moiré pattern, as shown in Fig. 2(d). Thus, the stringent requirement of camera positioning serves as a physical barrier, thwarting potential attackers from deciphering the Moiré code.

MoiréComm provides a number of benefits. To begin with, MoiréComm is a software-based approach that eliminates the need for external communication channels or supplementary hardware components. Furthermore, the decryption process hinge upon the relative position between the Moiré code and the capturing camera. Thus attackers are prevented from decrypting the Moiré codes due to the fact that they cannot occupy the same physical space as the would-be victim. Additionally, our proposed MoiréComm is built upon the existing QR code framework, and the entire encryption and decryption processes do not require changing the original QR code protocol; therefore, this makes it possible to easily deploy MoiréComm on the apps that require secure QR code communication, such as key exchanges and device paring.

We design three decryption methods for Moiré codes tailored to various screen-camera communication scenarios. In handheld scenarios, such as mobile payments where users and cashiers handhold smartphone screens and cameras, we offer a multi-frame decryption approach. This method employs traditional image processing algorithms to analyze and integrate the multiple captured Moiré codes with slighly hand tremors, reconstructing the original QR codes. This approach is universally applicable and particularly beneficial for embedded camera systems lacking deep learning support, although it requires more processing time. Considering that smartphones and scanning machines are typically equipped with powerful CPUs or GPUs that can support deep learning libraries like PyTorch. Based on this hardware foundation, we propose a decoding algorithm based on conditional generative adversarial networks (cGANs). This algorithm processes the captured Moiré codes with a pre-trained model for fast restoration of the original QR code, facilitating near-instantaneous decoding and mirroring the experience of scanning traditional QR codes. In stationary camera and screen setups, we propose a fast decoding technique using differential analysis. It involves displaying a reference image on the transmitter's screen, which is an encrypted Moiré pattern of a pure white base, and capturing it alongside the Moiré codes with the receiver's camera. The original QR code is retrieved by subtracting the reference image from the captured Moiré codes.

Comprehensive series of experiments are conducted to evaluate the MoiréComm for its effectiveness and robustness across various electronic screen, smartphone cameras, and Raspberry Pi cameras. Under handheld camera conditions, our proposed general multi-frame decryption method achieves a decoding accuracy rate exceeding 98.6% within an average of 13.2 frames, but with an average latency of 2.7 s. Furthermore, our cGAN-based decryption approach demonstrates an average decoding latency of just 0.02 s and an average success rate of 98.8% when using up to three frames captured near the target location. In scenarios with fixed camera and screen positions, our rapid decryption scheme attained a 99.4% success rate within two frames, with an average latency of approximately 0.4 s. The experiments also reveal that any unauthorized cameras distanced more than 10 cm away from the designated receiver location or angled beyond 20° experienced a decoding rate plummeting to zero, ensuring that potential attackers could not acquire useful information.

The contributions of this work include the following:

- We propose MoiréComm, including Moiré-based encryption and decryption techniques for improving the security of the screen-camera communication.
- We build a mathematical model to describe the Color Filter Array in the camera module for use in camouflaging spatial patterns via phase modulation and frequency modulation.
- We design three robust decryption schemes for the reconstruction of original QR codes from captured Moiré patterns in different application scenarios.

[1] A supplementary video demonstrates the organic unmasking process of an Moiré code: https://youtu.be/D10J7WCik8U. It should be noted that the scaling of Fig. 2(b) is tailored for this publication, and as such, it is not feasible to decrypt it using standard smartphone cameras.

- We implement the MoiréComm system in Android, iOS, and Raspbian (the embedding operating system), and conduct extensive experiments to assess the feasibility and limitations of the proposed MoiréComm. The source code is available: https://github.com/SolskyPan/mQRCode.

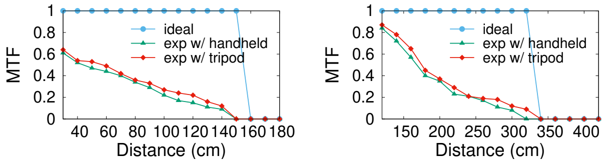
II. Threat Model

A. Mobile payment scenario

We envision a mobile payment scenario in which a user (hereafter referred to as the victim) intends to make a payment using a mobile payment application. During this process, the victim displays the payment QR code on their smartphone screen, which is then scanned by the cashier to complete the transaction. Simultaneously, an attacker attempts to covertly obtain the victim's QR code in order to commit fraud. The attacker may either approach the victim to secretly photograph the code using a smartphone or digital camera, or employ concealed cameras installed in the ceiling or wall outlets to surreptitiously capture the QR code. Since users often display the payment QR code before reaching the cashier, as documented in reported cases [7], there is a considerable likelihood that an adversary can successfully capture the code. To further understand how early users tend to show the payment QR code on the screen, we conducted both online and offline surveys in three cities in China, collecting responses from 150 users to investigate their habits regarding QR code payment while waiting in line. The details are shown below:

Questionnaire analysis: The questionnaire focused on three main questions: *(1) whether they are accustomed to using QR codes for payment; (2) whether they tend to open the QR code in advance before reaching the cashier; and (3) how many people are usually ahead of them in the queue when they open the QR code.* The results show that all participants use QR codes for payment. Among them, 73% open their payment QR code in advance while waiting in line. Specifically, 56% do so when there is only one person ahead, 36% when there are two people ahead, and 8% when three or more people are ahead. Overall, 73% of users prepare their QR code at least 30 seconds before it is scanned by the cashier.

According to the questionnaire results, the time available to the attacker is considered sufficient. Thus, once the victim's payment QR code is obtained, the attacker can immediately transmit it to a remote server to perform unauthorized operations. In addition, the attacker may interfere with the legitimate payment process by creating physical obstructions or employing social engineering tactics, such as engaging the cashier in conversation to delay the scanning process. After the unauthorized transaction is completed, the victim receives a payment notification; however, by that time, the attack has already succeeded and the unauthorized transaction has been finalized. Furthermore, as documented in several reports, victims often ignore such notifications or fail to recognize them as indicators of a security breach [7], [8].



(a) 18-105 mm lens (b) 70-200 mm lens

Fig. 3. Measured MTF of Nikon D7000 with two lenses.

B. Possible attacks to MoiréComm

The purpose of our proposed MoiréComm is to enhance the security of QR codes by leveraging the nonlinear properties of light and employing Moiré cryptography techniques to encrypt the original QR code image through pixel-level modulation. Authorized recipients can decode the original QR code information only within a designated physical space. We assume that the attacker is aware of the details of our encryption method, and discuss two possible strategies they might employ to compromise our system.

One possible attack method involves the attacker knowing the specific position required to photograph the victim's screen in order to successfully capture an Moiré code. The attacker could then further decode this image via our proposed decryption methods to obtain the victim's original payment QR code.

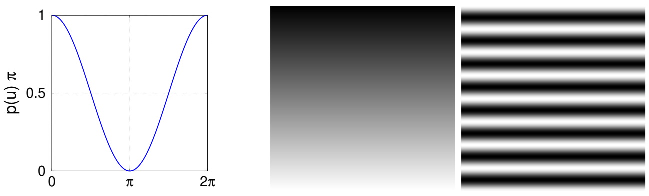
Another possible attack is that the attacker attempts to use an ultra-high-definition camera with a telephoto lens to capture, from a distance, every pixel of the Moiré code displayed on the victim's screen. Then, they can subsequently decode the original QR code. Specifically,detailed information about the spatial frequency and phase of the camera's Bayer CFA is publicly available. After thoroughly studying our work, the attacker can use the encryption algorithm to determine the specific value (*i.e.*, white or black) of each pixel in the encrypted QR code. Through reverse engineering, it is then possible to infer the information of the original QR code based on these pixel values.

C. Countermeasures again attacks

For the first attack method, it would require the attacker to be in close proximity to the victim, for example, within 50 cm of the victim's smartphone screen and at a viewing offset angle of less than 6 degrees. Such behavior would be highly conspicuous and therefore difficult to execute in practice. Thus, we argue that this attack strategy is unlikely to succeed.

For the second attack method, it is theoretically feasible to obtain the raw QR code if the attacker can capture every pixel of the Moiré codes from the victim's screen. However, how feasible is it for the attacker to obtain pixel-level information in practice? Next, we analyze the feasibility for an attacker to obtain pixel-level information from a screen at a distance.

According to the ideal pinhole imaging principle, the spatial resolution is determined by the pixel pitch of the display and the camera as well as the camera focal length. However, in practice, spatial resolution is largely impaired by lens distortion and aliasing [11]. One well-known metric used to quantify spatial resolution is the Modulation Transfer Function (MTF). The modulation of an image represents its contrast in spatial domain, as follows: $M = \frac{S_{max}-S_{min}}{S_{max}+S_{min}}$, where S_{max} and S_{min} represent the maximal and minimal pixel values



(a) $p(u) = 0.5 + 0.5cos(u)$ (b) $\phi_1(x, y) = x$, $p(\phi_1(x, y))$

Fig. 4. Examples of periodic ($p(\cdot)$) and phase ($\phi(\cdot)$) functions in a 2D spatial structure

within an image. A higher M is indicative of higher contrast. MTF is used to define the modulation ratio between a captured image and the displayed image: $MTF = \frac{M_{cap}}{M_{disp}}$, where M_{cap} and M_{disp} respectively represent the modulation of the captured image and the displayed image. We usually use a displayed image with $M_{disp} = 1$, such that MTF falls within a range between 0 and 1. The MTF is proportional to the spatial resolution. We can set a cutoff threshold for MTF. When the MTF value is below the threshold, there is no way to differentiate among black and white pixels. Fig. 3 shows the MTF of the Nikon D7000 with an AF-S NIKKOR $18 \sim 105\ mm$ lens and an AF-S NIKKOR $70 \sim 200\ mm$ lens. When we set the cutoff threshold to 0.3 [12], the camera with 105 mm focal length is unable to differentiate individual pixels when the camera is farther than $1.08\ m$. We can notice the measured distance is much smaller than $1.40\ cm$ computed using the ideal pinhole imaging principle. When the camera is held in hand (which is a more realistic case for the attacker), MTF is further decreased by 3.92%. When a telephoto lens with the 200 mm focus is used, the maximal attack distance is increased to $2.1\ m$. However, the size of the lens is also increased to $88.5\ mm \times 202.5\ mm$ (diameter and length) which makes it even harder to disguise. Although lenses with longer focal lengths (*e.g.*, 800 mm [13]) are available in the market, their sizes and prices make them hard to be used in the attacks. Overall, it would be reasonable to conclude that MoiréComm greatly reduces the risks involved in leaking information via QR code.

In contemporary flagship smartphones, the equivalent focal length of telephoto lenses typically ranges from 120 to 240 mm[14]. To examine the capability of capturing encrypted QR images, we selected the Xiaomi 13Ultra, which is equipped with a telephoto lens with a 240 mm equivalent focal length (in 10× optical zoom mode), as the attacker mobile device. Our findings revealed two main issues: (1) focusing on nearby objects in telephoto mode proved challenging; and (2) even at successful focus, the effective distance for the Xiaomi 13Ultra to capture pixel-level information from a screen was limited to $0.27\ m$. This primarily arises because, although the mobile camera features a 240 mm lens, its camera sensor, the IMX858 [15], has a pixel size only a quarter of the Nikon D7000 [16], greatly restricting the smartphone's ability to capture pixel details. Consequently, the effective range for capturing screen pixel information overlaps with the Moiré pattern receiving area. Therefore, ensuring that smartphones are absent within the receiving area can effectively prevent mobile eavesdropping.

After analyzing the potential attack methods, we recommend the following countermeasures for victims: (1) Ensure that only the cashier's camera is within the acceptable range of the payment device; (2) Make sure there are no concealed smartphones with telephoto lenses within one meter, and no telephoto digital cameras within two meters, of the payment device. Additionally, the victim can orient the QR code screen towards the ceiling when presenting it to the cashier, further reducing the likelihood of being surreptitiously photographed at the pixel level.

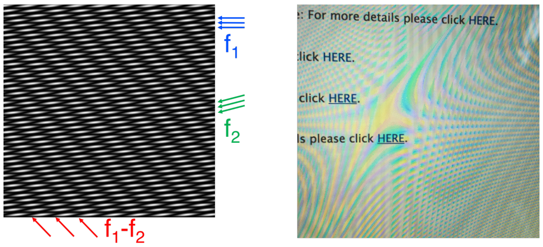


Fig. 5. Left is nonlinear optical interaction of two patterns with frequency f_1 and f_2; right is nonlinear optical interaction when taking a picture of a display using a camera

III. Background

A. Spatial Frequency

Spatial frequency pertains to the recurring nature of a structure's pattern as it periodically extends through space. Our work is centered on a curvilinear pattern within a two-dimensional (2D) spatial framework, *i.e.*, image. To characterize this 2D pattern, we employ both a frequency and a phase term, as delineated below:

$$m(x, y) = p(\phi(x, y)) \tag{1}$$

where $m(x, y)$ denotes the magnitude at a specific 2D coordinate (x, y), *i.e.*, the color of an image, $p(\cdot)$ is a periodic function representing the the pattern's frequency, while $\phi(x, y)$ is a phase function representing the pattern's angle. For example, Fig. 4(a) shows the periodic function using a cosine wave with a frequency of $1/2\pi$. By setting the phase function to $\phi(x, y) = x$, a repetitive horizontal striped pattern emerges, as depicted in Fig. 4(b).

B. Nonlinearity of Spatial Frequency

The superimposition of two unique spatial patterns results in nonlinear optical interactions that manifest as an additional visible layer atop the original patterns, known as a Moiré pattern. In the process of image overlay, the multiplication of pixel values is employed rather than addition to more accurately reflect the physical phenomena of light transmission and absorption. In the context of gray-scale images, the value of each pixel is indicative of the level of light reflectance, with a value of 0 representing complete light absorption (black) and a value of 1 indicating full light reflection (white). The multiplication of corresponding pixel values from two images yields a composite pixel value that signifies the combined effect on light reflectance. This principle is also applicable to color images and is applied to each color channel, *i.e.*, red, green, and blue, independently. Therefore, if we denote m

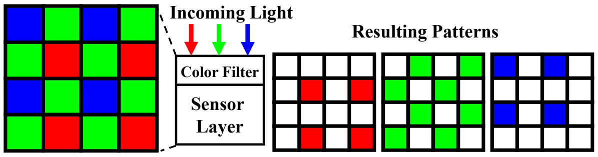


Fig. 6. Profile of sensor with the Bayer arrangement of color filters

as the resultant image from the superposition of two layers, m_1 and m_2, we can determine m through the following calculation:

$$m(x, y) = m_1(x, y) \times m_2(x, y) \tag{2}$$

The multiplicative model produces nonlinearities in the spatial frequency. For example, when m_1 and m_2 use cosine functions with frequency f_1 and f_2 as periodic functions:

$$\begin{aligned} m =& m_1 \times m_2 \\ &= (a_1 + b_1 cos(2\pi f_1 t)) \times (a_2 + b_2 cos(2\pi f_2 t)) \\ &= a_1 a_2 + a_1 b_2 cos(2\pi f_2 t) + a_2 b_1 cos(2\pi f_1 t) \\ &\quad + b_1 b_2 cos(2\pi (f_1 + f_2) t) + b_1 b_2 cos(2\pi (f_1 - f_2) t) \end{aligned} \tag{3}$$

and the combination result includes two additional frequencies $(f_1 + f_2)$ and $(f_1 - f_2)$. While human eyes are more sensitive to low frequency signals; therefore, frequency $(f_1 - f_2)$ is easier to observe, as shown in the left of Fig. 5.

Taking into account the generality of patterns, when m_1 and m_2 are represented as curvilinear patterns, where $m_1 = p_1(\phi_1(x, y))$ and $m_2 = p_2(\phi_2(x, y))$, the convolution theorem [17] can be applied to perform spectral analysis on their composite pattern m:

$$M(x, y) = M_1(x, y) \otimes M_2(x, y) \tag{4}$$

where M, M_1, and M_2 represent the Fourier Transform of m, m_1 and m_2, respectively; and the $\otimes$ operator represents the 2D convolution. According to Moiré theorem [18], the periodic function and phase function are independent and can therefore be computed separately. Let m_{nl} represent the evident nonlinear component resulting from the superposition of m_1 and m_2 with frequencies $(f_1 - f_2)$. Due to the fact that m_{nl} is also a curvilinear pattern, $m_{nl} = p_{nl}(\phi_{nl}(x, y))$ is in accordance with Eq. 1. We then decompose Eq. 4 and compute its periodic function $p_{nl}(u)$ and phase function $\phi_{nl}(x, y)$ as follows:

$$\begin{aligned} p_{nl}(u) &= IFFT(FFT(p_1(u)) \cdot FFT(p_2(-u))) \\ \phi_{nl}(x, y) &= \phi_1(x, y) - \phi_2(x, y) \end{aligned} \tag{5}$$

where $FFT(T)$ and $IFFT(T)$ are the Fourier Transform and Inverse Fourier Transform of input T, respectively; and $p_1(u)$, $p_2(u)$, $\phi_1(x, y)$, and $\phi_2(x, y)$ are the corresponding periodic and phase functions for m_1 and m_2.

C. Nonlinearity in Camera Systems

Cameras, as nonlinear systems, are prone to the generation of Moiré patterns when capturing images of spatial patterns. An illustrative case is depicted on the right side of Fig. 5, where an image captured by a camera displays a set of curving lines that are overlaid onto the screen pattern. This nonlinearity stems from the color filter array, which consists of a mosaic of minuscule color filters positioned above the pixel CMOS sensors to facilitate the acquisition of color data. Among these, the Bayer CFA is the most frequently utilized [19], especially in mobile device-embedded cameras. As demonstrated in Fig. 6, the Bayer CFA captures light intensity levels for one red, two green, and one blue channel within every 4×4 matrix.

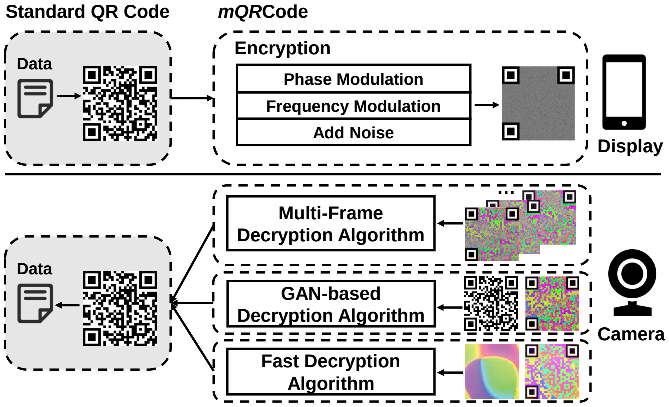


Fig. 7. System overview of our proposed MoiréComm, incorporating Moiré-based encryption and decryption methodologies into the existing camera-screen communication framework to enhance the security

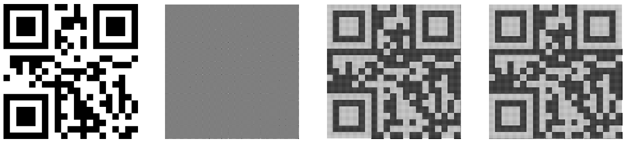
During the process of photographing a screen, the pixel array of the LCD or OLED screen creates a spatial pattern with frequency f_1 when projected onto the camera, and the CFA in the camera itself imposes an additional frequency f_2. By carefully positioning the camera at a specific distance and angle relative to the screen, the resulting frequency disparity $(f_1 - f_2)$ can be adjusted to fall within an observable frequency band. This leads to a nonlinear optical effect that manifests as a rippled pattern in the captured image, as shown on the right side of Fig. 5. In Sec. IV, the forthcoming discussion will elaborate on the method by which the MoiréComm harnesses this nonlinear optical interplay between the camera and the screen to encrypt QR codes as camouflage images, thereby enhancing the security of screen-camera communications.

IV. MoiréComm Design

A. System overview

The primary aim in designing MoiréComm is to enhance the security of the camera-screen communication system, *e.g.*, QR code. The architecture of MoiréComm is depicted in Fig. 7, and comprises two main components: a transmitter (*e.g.*, a smartphone screen) and a receiver (*e.g.*, a smartphone camera). As the transmitter produces a conventional QR code, MoiréComm introduces both phase modulation (detailed in Sec.IV-B2) and frequency modulation (described in Sec. IV-B3) techniques to encrypt the QR code. To further secure the encryption and manage phase discontinuities, the encrypted QR code (denoted as Moiré code) is masked with strategically placed noise. This process generates an Moiré code which, when displayed, offers an elevated level of security.

On the receiver side, we provide three distinct decryption methodologies tailored to suit various deployment contexts.



(a) QR Code (b) Green filter (c) Red filter (d) Blue filter

Fig. 8. Encrypt a QR code image using color filters in Bayer CFA

It is important to position the camera at a precise distance and angle in relation to the Moiré code for optimal functionality. For circumstances where the devices are mobile, with either the transmitter screen or receiver camera being handheld by users, we design a conventional multi-frame based decryption approach as a comprehensive solution (detailed in Sec. IV-C2). Additionally, for those receiving devices with the computational power to execute deep learning libraries such as PyTorch, we develop a GAN-based decryption technique aimed at diminishing decryption time (elaborated in Sec. IV-C3). In the case of stationary scanners, which are a common fixture in retail settings, the screen showcasing the Moiré code is situated under the scanner to initiate the fast decryption process (as outlined in Sec. IV-C4). Once the Moiré code is successfully restored to its original form, we utilize a standard QR code decoder to retrieve the data embedded within the code, ensuring seamless integration with existing QR code reading technology while also providing a secure transmission channel.

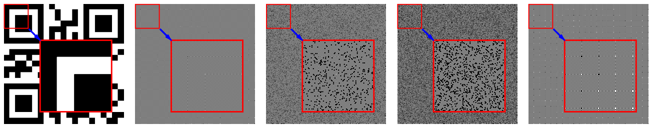
B. Encryption

MoiréComm harnesses the nonlinear optical interaction that occurs between the CFA and the encryption pattern overlaying the QR code. As established by Eq. 2, we assume a general scenario where the spatial configuration of the CFA is denoted as $m_{cfa}(x, y)$, and the original QR code—which is synonymous with the decrypted image—is represented by $m_{dec}(x, y)$. The encryption objective is to ascertain the coded QR code image, $m_{enc}(x, y)$, in such a way that the relationship $m_{dec}(x, y) = m_{cfa}(x, y) \cdot m_{enc}(x, y)$ holds true.

1) Color Filter Array Model: We first model the CFA in the camera using the equation $m_{cfa}(x, y) = p_{cfa}(\phi_{cfa}(x, y))$. In the Bayer filter (see Fig.6), green filters are located on the diagonals of each 2×2 cell, while blue and red filters occupy the remaining positions. Instead of modeling the spatial frequency characteristics of all three color channels, MoiréComm utilizes only the green filters, so we model only the spatial frequency of the green filter. There are two reasons for this: (1) QR codes contain only black and white blocks, so one color filter is sufficient for decryption. (2) the green filter in the 2×2 array is symmetric, which prevents abrupt phase changes from being distinguished by the human eye or the camera. Thus, the green filter is modeled as follows:

$$\begin{aligned} m_{cfa}(x, y) &= p_{cfa}(\phi_{cfa}(x, y)) \\ p_{cfa}(u) &= 0.5 + 0.5cos(2\pi u) \\ \phi_{cfa}(x, y) &= ((x + y) \bmod 2)/2 \end{aligned} \tag{6}$$

where $m_{cfa}(x, y)$ represents the color reception of the green filter at coordinate (x, y) on the image sensor, $p_{cfa}(u)$ represents the periodic function, and $\phi_{cfa}(x, y)$ represents the phase function. Since x and y range from 1 to the image height/width, values of the phase function become 0 or 0.5. After application of the periodic function, $m_{cfa}(x, y)$ becomes 1 (*i.e.*, the CMOS sensor receives all light on the green filter) in the diagonal grids in each 2×2 array or 0 (*i.e.*, the CMOS sensor receives no light).



(a) QR code (b) Moiré code (c) 10% dot (d) 20% dot (e) Lines

Fig. 9. Addition of noise or camouflaging lines to mitigate observable lines caused by abrupt phase changes

2) Phase Modulation: Our goal is to ensure that the information in the $m_{dec}(x, y)$ is as the same as to that of the raw QR code image. Here, we assume that each block of the QR code represents a DC value, *i.e.*, its spatial frequency is zero. Thus, based on the properties of nonlinearity, we need to make the spatial frequencies of the encoder (p_{enc}) and the CFA (p_{cfa}) as the same. In this way, after the multiplication of the two spatial frequency functions, one of the nonlinear components (*i.e.*, lower frequency component), *i.e.*, Moiré pattern, can convey the information of each block in the QR code [20]. Thus, we let $p_{enc}(u)$ equal $p_{cfa}(u)$:

$$p_{enc}(u) = 0.5 + 0.5cos(2\pi u) \tag{7}$$

QR code images consist of two types of blocks: black and white. We apply phase modulation to map two types of blocks to different phases, to make them distinguished after the nonlinear multiplication. Due to the fact that m_{dec} (*i.e.*, original QR code) and m_{cfa} (modeled in Sec. IV-B1) are known, combining Eq. 2 and 5 we have:

$$\begin{aligned} m_{dec}(x, y) &= p_{dec}(\phi_{dec}(x, y)) \\ &= p_{dec}(\phi_{cfa}(x, y) - \phi_{enc}(x, y)) \\ \Longrightarrow \phi_{enc}(x, y) &= \phi_{cfa}(x, y) - p_{dec}^{-1}(m_{dec}(x, y)) + 2k\pi \end{aligned}$$

where $k \in \mathbb{Z}$, and p_{dec}^{-1} represents the inverse function of p_{dec}, which maps intensity values to the corresponding phases. The $2k\pi$ term has no impact on the encrypted image m_{enc} because we use cosine as the periodic function (Eq. 7).

By using only the phase modulation method (the same method in [20]), where the phase information of the black and white colors is directly inverted, we observe that phase transitions at the interface of black and white segments may lead to the appearance of conspicuous horizontal or vertical streaks within the encrypted image, as depicted in Fig.9(b).

To address this issue, we introduce different types of disruptive noise. Specifically, the results with the introduced dot noise are illustrated in Fig.9(c) and Fig.9(d), while the result with obfuscating stripes is shown in Fig. 9(e). Empirically, adding 10% noise is sufficient for Moiré codes displayed on a smartphone screen with extremely high camouflage. By incorporating noise into the phase modulation, we can effectively mitigate the problem of phase discontinuity. This encoding method is more secure than that in [20], as phase

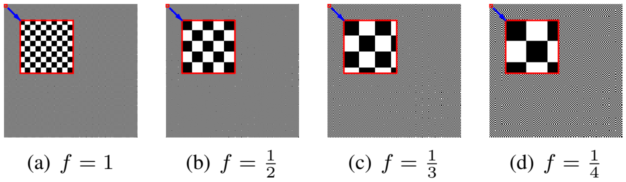

(a) $f = 1$ (b) $f = \frac{1}{2}$ (c) $f = \frac{1}{3}$ (d) $f = \frac{1}{4}$

Fig. 10. Example of using various frequencies f to communicate at various distances

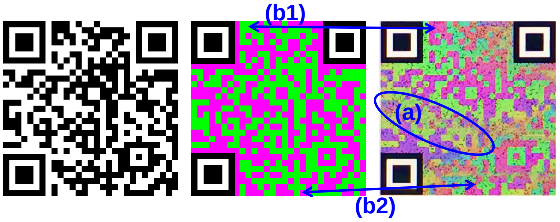


Fig. 11. Challenges of decrypting mQR codes. The area (a) enclosed by the blue circle shows the blur in the Moiré code captured by the camera; Parts (b1) and (b2) illustrate the inversion phenomenon compared to the ground truth of the Moiré code

discontinuity could potentially lead to the leakage of the original image information from the encryption.

3) Frequency Modulation: The numerous applications to which MoiréComm could be applied may require that camera capture the pictures at different distances from the screen. Recall that when a camera captures an encrypted QR code on a screen and it undergoes a nonlinear multiplication operation, the spatial frequency of the Moiré pattern is $(f_1 - f_2)$. Here, f_1 is the spatial frequency of the camouflage pattern projected onto the CMOS through the lens, and f_2 is the spatial frequency of the CFA. Our previous premise for phase modulation was that the Moiré pattern becomes more pronounced when f_1 is close to f_2. Therefore, the distance between the lens and the screen significantly affects the parameter f_1. We define the Moiré distance (D_m) of the screen-camera pair as the camera capture the Moiré pattern from the screen when the screen modulates the image by one pixel.

According to camera pinhole theory [21], the size of an object projected onto a camera sensor is inversely proportional to the distance between the object and the camera sensor as: $S_{cam} = \frac{S_{object} \times L_{focal}}{D}$, where S_{cam} is the object size in the captured image, S_{object} is the object size in the real world, L_{focal} is the focal length of the camera, and D is the distance between the camera and the object. Although the pinhole camera model does not account for lens distortion, we can derive an estimate with regard to Moiré communication distance, *i.e.*, D_m, for a given camera and screen parameters.

Notably, when the receiver's camera captures the Moiré codes, increasing the image size using digital zoom does not affect the communication distance D_m, since the optical focal length remains unchanged. However, if optical zoom is used to enlarge the image size, D_m will increase as the focal length increases. Fortunately, receiver cameras, such as cash register scanners or users' mobile phones, typically use a fixed focal length or only employ digital zoom during close-range QR code communication. In such cases, the predetermined communication range remains unaffected.

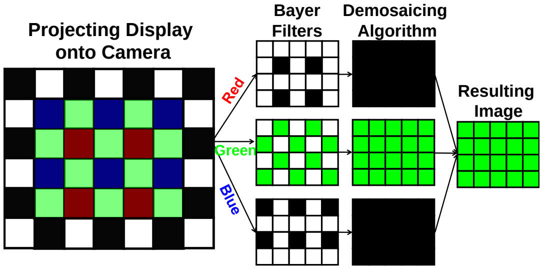
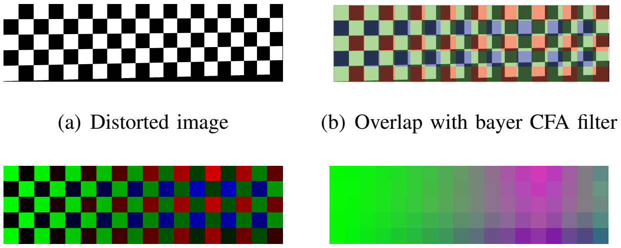


Fig. 12. Analysis illustrating Moiré patterns in encrypted images: simulation-based alignment results in the green filter


(a) Distorted image (b) Overlap with bayer CFA filter

(c) Bayer pattern image (d) After interpolation

Fig. 13. Simulation-based analysis illustrating phase inversion

To enable support for various communication ranges, we extend Eq. 6 to modulate the frequency (f) of the generated spatial patterns, as follows:

$$\begin{aligned} m_{cfa}(x, y) &= p_{cfa}(\phi_{cfa}(x, y)) \\ p_{cfa}(u) &= 0.5 + 0.5cos(2\pi u) \\ \phi_{cfa}(x, y) &= ((\lceil xf \rceil + \lceil yf \rceil) \text{mod} 2)/2 \end{aligned} \tag{8}$$

where $f \in \{1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, ...\}$. Fig. 10 shows examples of Moiré codes in which frequency modulation is applied to adjust the communication distance with 1, 2, 3, and 4 times the D_m. One drawback we can see in Fig. 10 is that when we use a lower frequency, the boundary of QR code blocks becomes more evident due to the abrupt phase change. We address this issue with the introduced noise to obscure the distinct boundary information, the same method described in Sec. IV-B2.

C. Decryption

Decrypting the Moiré code requires that the user holds the camera in a designated position, whereupon the Moiré effect reveals the original QR code. However, using this image directly to reconstruct the original QR code is challenging, due to the existence of blurred portions and phase inversion. Examples of blurred portions are presented in Fig. 11(a). Due to the effects of phase inversion, blocks with the same color in the original QR code (*i.e.*, blocks which are modeled using the same phase in Moiré code) may end up exhibiting different colors in the Moiré pattern, as shown in Fig. 11(b). Specifically, (b1) shows black blocks of the original QR code mapped to purple in the captured picture, whereas (b2) shows black blocks mapped to green. Before describing the decryption methods, we first propose a simulation-based analysis to better understand the phenomena.

1) Challenges: Phase Inversion and Blur: Phase inversion refers to the phenomenon in which blocks of the same color

This article has been accepted for publication in IEEE Transactions on Dependable and Secure Computing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TDSC.2025.3598300

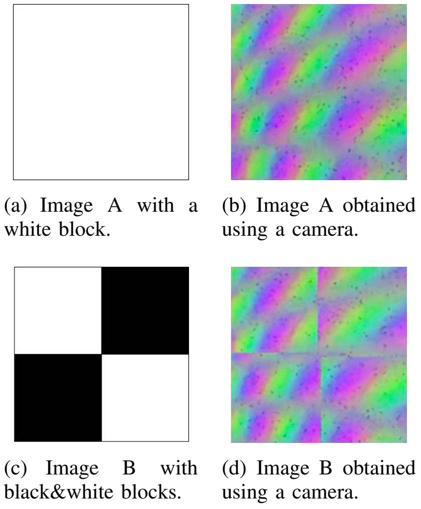

(a) Image A with a white block.
(b) Image A obtained using a camera.
(c) Image B with black&white blocks.
(d) Image B obtained using a camera.

Fig. 14. Illustration of difference in Moiré patterns in images A and B under the same display and camera conditions

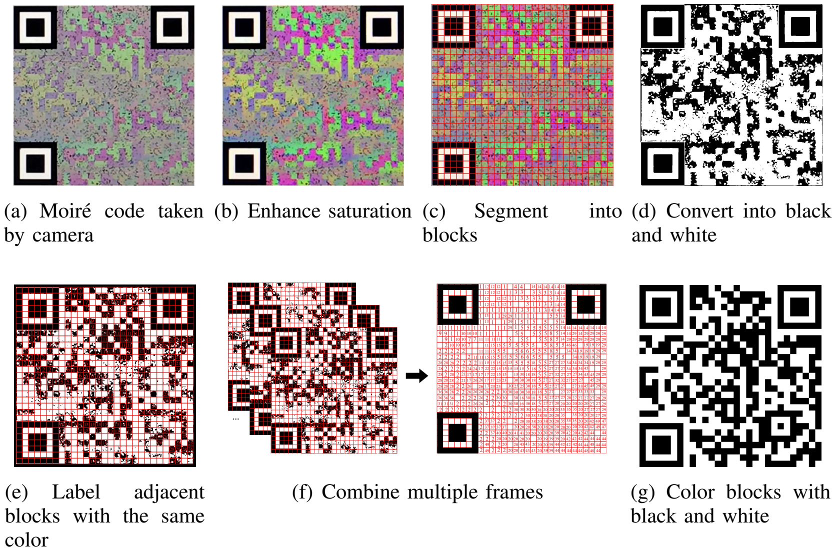

(a) Moiré code taken by camera
(b) Enhance saturation
(c) Segment into blocks
(d) Convert into black and white
(e) Label adjacent blocks with the same color
(f) Combine multiple frames
(g) Color blocks with black and white

Fig. 15. Traditional multi-frame decryption scheme applied in handheld screen and camera scenarios

in the original QR code (*i.e.*, blocks modeled using the same phase in Moiré code) end up exhibiting different colors in the resulting Moiré pattern. We adopted simulation-based analysis to illustrate how phase inversion occurs. The bottom left part of Fig. 12 shows part of an Moiré code in which each block represents a pixel, and these pixels are either black or white, based on the phase modulation scheme (Eq. IV-B2). The top left part of Fig. 12 presents the Bayer CFA of a camera in which each CMOS photosensor captures red, green, or blue light.

In an ideal scenario, placing a camera with no lens distortion precisely in the designated position will allow the perfect alignment of pixels on the display, as shown in the left part of Fig. 12. The raw output from a photosensor is referred to as a Bayer pattern image. Obtaining a full-color image requires a variety of demosaicing algorithms [22], [23], [24], [25], which interpolate red, green, and blue values for each pixel. In the process, the three channels are merged, resulting in the creation of an image that is perceived by the human eye in its intended form. For instance, within this specific example, the green channel alone is responsible for capturing the screen's high-intensity light, leading to the exclusive visibility of the green color. If the screen is displaced by a single pixel, it is the red and blue channels that predominantly receive the high-intensity light, thus presenting us with a purple hue due to their fusion. Under such an optimal condition, the captured image of the Moiré code exhibits a uniform phase correspondence throughout. This means that, within the entire image, the black squares of the original QR code consistently correspond to either green or purple, as exemplified by the central illustration in Fig. 11.

Nonetheless, when lens distortion is present or the camera is not accurately aligned with the intended position, the image projected onto the camera's sensor can become distorted. As demonstrated in Fig. 13, with the camera tilted by 1°, the screen's image is cast onto the photosensor array with some distortion. Subsequently, we process the Bayer pattern of the projected image, interpolate to generate the full-color representation, and merge the separate channels to replicate the image as it would be captured by the camera. The resulting Moiré pattern exhibits a green hue on the left side and a purple one on the right, vividly demonstrating the occurrence of phase inversion through the simulation. Utilizing the encryption method we propose, a white image (as shown in Fig. 14(a)) is encrypted, resulting in a Moiré pattern (depicted in Fig. 14(b)). It is observed that despite the encrypted image retaining identical phase information, its associated Moiré pattern displays multiple colors. This exemplifies the complexity involved in devising a robust decryption strategy.

2) Traditional Multi-Frame Decryption Scheme: Our preliminary analysis has revealed that distortions from the camera lens and inaccuracies in camera positioning can lead to phase inversion, complicating the decryption process of an Moiré code when using a single snapshot. To overcome these challenges, we firstly design the MoiréComm decryption algorithm that leverages multiple consecutive frames from a video. This algorithm is based on the observation that when a user holds the camera, there is inevitably a certain amount of camera shake (usually less than $2 \sim 3mm$) [26]. The minute motions between frames result in varying degrees of blur and phase inversion across different areas of the Moiré code images. By analyzing the disparities across these frames, it becomes possible to reconstruct the original QR code. The multi-frame decryption algorithm that we develop is comprised of the following steps:

Enhancing color saturation: Fig. 15(a) presents a picture of an Moiré code captured by a smartphone camera. We first enhance the color saturation to improve contrast among green, red, blue, and resulting image is shown in Fig. 15(b).

Segmentation: A standard QR code is equipped with three distinct positioning markers situated at predetermined locations, which facilitate the detection of a QR code by a reader, ascertain the version of the QR code, and align the image accordingly. The MoiréComm approach retains these critical locators intact. Indeed, we employ these positioning markers

This article has been accepted for publication in IEEE Transactions on Dependable and Secure Computing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TDSC.2025.3598300

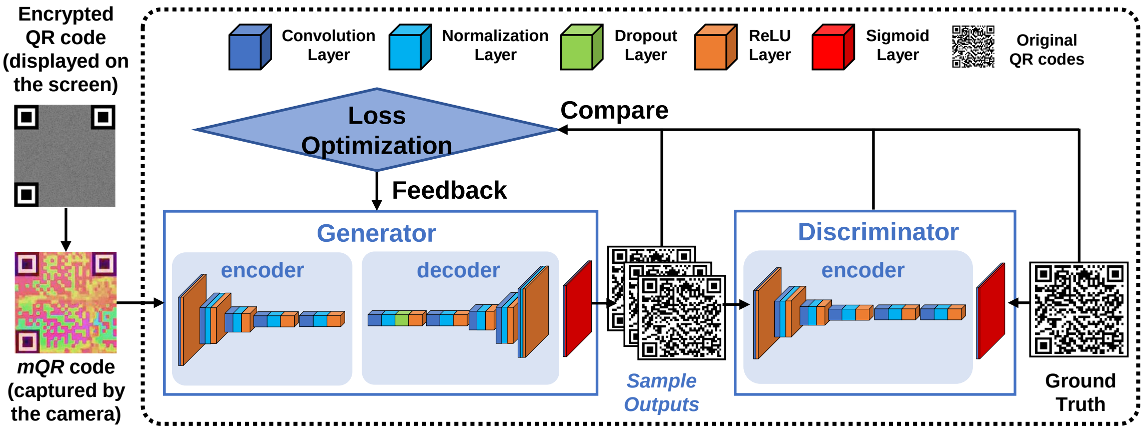


Fig. 16. Conditional GAN-based decryption scheme applied in handheld screen and camera scenarios.

as reference points to construct lines for perspective correction, transforming skewed squares back into their proper orthogonal shape. Subsequently, the dimensions of these locator marks are utilized to ascertain the dimensions of each individual block within the QR code. An illustration of a segmented Moiré code is depicted in Fig. 15(c).

Conversion to black and white: In Sec. IV-B, we describe the decryption of QR codes by modeling green filters and modulating the phase function to generate Moiré codes. QR code blocks with the different colors are assigned different phases, resulting in either green or purple separation in the Moiré pattern, as shown in Fig. 15(b). However, phase inversion alters the color mapping in the spatial domain. To reliably identify blocks with the same phase, we separate green from purple by thresholding the green channel and converting the image to black and white. In other words, when the green intensity of a pixel is higher than a given threshold, then the pixel is changed to white; otherwise, the pixel is changed to black, as shown in Fig. 15(d).

Classification of blocks: After each block is changed to black or white, noise can result in both black and white pixels in a given block. Thus, in classifying each block as black or white, we calculate the proportion value c of black pixels in each block. When c is larger than a given threshold (in this case 0.8), the block is classified as black; otherwise, the block is classified as white.

Labeling adjacent blocks with the same color: Two adjacent black blocks probably have the same phase in Moiré code. Thus, we loop through all of the black blocks and label them using an index. Adjacent blocks that are both black are labeled using the same index, as shown in Fig. 15(e).

Combining multiple frames: The above steps are repeated for each incoming frame. The labels from the new frame are then combined with existing labels from previous frames as follows: If a block does not have an existing label or is assigned a label in the new frame, then the block is assigned a new label. If a block has an existing label $index_{old}$ and is assigned another label $index_{new}$ in the new frame, then we search among existing frames for blocks with label $index_{old}$ and blocks with label $index_{new}$ in the new frame, and assign them a new label. We continue combining new frames until either all of the blocks are labeled or all of the blocks surrounding an unlabeled block are labeled. An example is presented in Fig. 15(f).

Coloring blocks: Each block is then colored black or white in accordance with the labels. The colors of the locator marks are known; therefore, we begin by coloring their neighbors. The rules for color blocks are as follows: i) If two adjacent blocks have the same label, then they are drawn using the same color, and ii) if two adjacent blocks have different labels, then they are drawn using different colors. The original QR code is then recovered after all of the blocks have been colored, as shown in Fig. 15(g).

3) Conditional GAN-based Decryption Scheme: Traditional multi-frame decryption scheme can work well when the camera is handheld to capture the Moiré codes and can be deployed on almost any of IoT devices that have CPU or single-chip microcomputer. However, this method has non-negligible decryption latency due to the need for pixel-level and block-level cyclic processing of multiple frames. Considering that current smartphones have powerful CPUs and/or GPUs to support libraries for deep learning (*e.g.*, PyTorch Mobile [27]), here, we realize a deep learning-based approach to realize efficiently decrypt Moiré codes in the handheld case.

Unlike Convolutional Neural Networks (CNNs), which typically rely on predefined loss functions to optimize network parameters, GANs utilize a dynamic adversarial framework that allows for flexible generation of output data based on given conditions [28], [29]. This adversarial setup, comprising a generator and a discriminator, enables conditional GANs (cGANs) to learn complex data distributions and produce high-quality samples that align closely with the input conditions. Inspired by the *Pix2pix* model [30], a cGAN model that solves the image-to-image mapping problem, we propose a cGAN-based decryption scheme to translate Moiré codes into original QR code images. An overview is shown in Fig.16. We first formulate the cGAN-based adversarial loss as follows:

$$\begin{aligned}\mathcal{L}_{\mathcal{GAN}}(D,G) = \mathbb{E}_{x\sim p_{data}(x)}[logD(x)] \\ +\mathbb{E}_{z\sim p_z(z)}[log(1-D(G(z)))]\end{aligned} \tag{9}$$

where G is the generator function, D is the discriminator function, x is the training data being Moiré codes (Moiré pattern) images taken by a camera, z is the target data being the original QR code images. To achieve the desired translation result, the generator G and the discriminator D need to collaborate and compete with each other. From the loss function of cGAN, we observe its primary mechanism: enhancing the quality of generated images through adversarial training between the generator and the discriminator. This competition encourages the generator to produce results consistent with the specific target image based on the input image because the discriminator gradually becomes more adept at discerning the authenticity of the generated images.

To ensure that the generated images are not only realistic but also consistent in content with the target images, pixel-level losses (such as L1 loss) can be used. We define the content loss as follows:

$$L_{content}(G) = \mathbb{E}_{x,z}[||G(x) - z||_1] \tag{10}$$

The overall loss function is a weighted sum of these components:

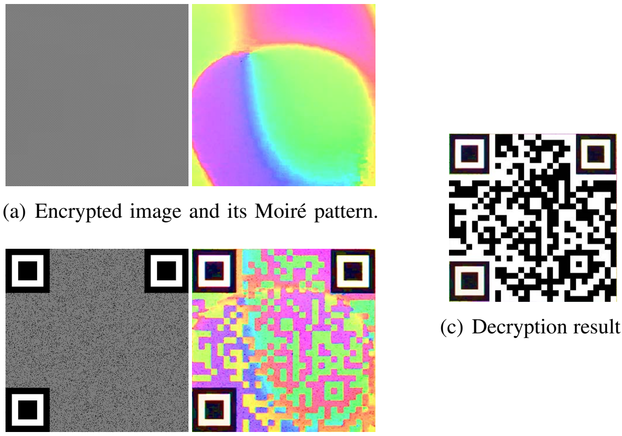
$$L_{total}(G, D) = \lambda_{GAN} L_{GAN}(G, D) + \lambda_{content} L_{content}(G) \tag{11}$$

Here, λ_{GAN} and $\lambda_{content}$ are hyperparameters that balance the contributions of each term. The complete loss function of cGAN inherently equips the encoder to generate the ground truth QR code images with the captured Moiré code as the input. Thus, it can provide the desired functionalities: 1. deblurring; 2. correcting phase inversion; 3. converting to the black-and-white style.

Generator is a network of “encoder-decoder” architecture. The encoder has one *Conv-ReLU* layer with 64 4×4 spatial filters and four *Conv-BN-ReLU* layers with [128,256,512,512] 4×4 spatial filters where Convolution downsamples by a factor of 2 and *ReLU* is leaky with slope 0.2, whereas the decoder has one *Conv-BN-Dropout-ReLU* layer with a dropout rate of 50% and four *Conv-BatchNorm-ReLU* layers where *Conv* upsamples by a factor of 2 and *ReLU* is not leaky. After the last layer in the decoder, an eventual convolution is applied to map to RGB 3 channels, followed by a *sigmoid* function.

Discriminator is a network of “encoder” architecture composed of one *Conv-ReLU* layer with 64 4×4 spatial filters and five *Conv-BN-ReLU* layers with [128,256,512,512,512] 4×4 spatial filters where *Conv* downsamples by a factor of 2 and *ReLU* is leaky with slope 0.2. After the last layer, an eventual convolution is applied to map to a 1D output, followed by a *sigmoid* function.

In our cGAN model, the generator creates sample outputs based on the input Moiré codes, while the discriminator determines if these generated images closely resemble the original QR code images. The output from the discriminator, along with the loss values, is sent to the optimizer, which adjusts the weights of both the discriminator and the generator accordingly. Ultimately, the trained generator can produce images that match the original QR code images, indicating that the cGAN-based decryption model has successfully learned the mapping function from the captured Moiré codes to the original QR code images. This encoder function can be regarded as our decryption scheme.



(a) Encrypted image and its Moiré pattern.

(b) Moiré code and its Moiré pattern.

(c) Decryption result.

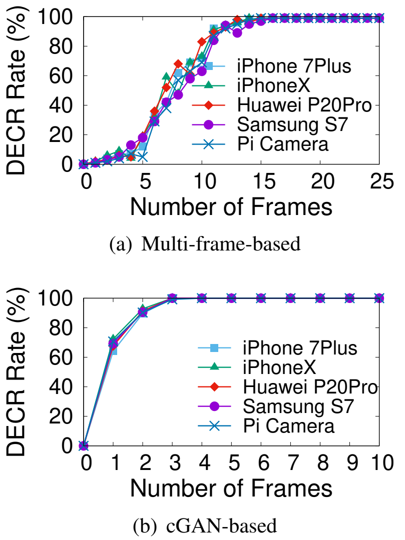
Fig. 17. Fast decryption applied in fixed screen-camera scenarios

4) Fast Decryption Scheme: The above proposed two decryption schemes are effective when users hold the camera in hand. Considering that many QR code scanners used in stores nowaday are fixed on a table [31], therefore, we sought to develop a fast decryption scheme for the scenario with a fixed scanner.

For the sake of illustration, we present the following simple experiment. We first use the proposed encryption scheme to encrypt a white image (as shown in Fig. 14(a)) and its Moiré pattern (Fig. 14(b)). We then encrypt another image with 2×2 black and white blocks, the Moiré pattern of which is shown in Fig. 14(d). The sizes of the two images are the same, and the display and camera are fixed in set positions. Clearly, the shape of the Moiré patterns from two images are similar; however, the colors located at black blocks in the second image are inverted. This suggests that when the camera and the display configuration are unchanged and the Moiré pattern of the encrypted white image is known, then we can predict the Moiré pattern of any Moiré codes in a given position simply by inverting the green and purple located at black blocks.

This observation is used to guide the design of our fast MoiréComm decryption scheme. We employ a QR code scanner and placed a smartphone displaying an Moiré code on the table to facilitate scanning. The phones alternatively display an encrypted white image and an Moiré code at $10fps$. The camera is configured to use a fixed focal length while recording video at a frame rate of $30fps$. Once the scanner captures from the phone a frame with the encrypted white image and a frame with the Moiré code, the two frames then undergo processing to reconstruct the original QR code. In details, we utilize the captured the Moiré pattern of the encrypted white image as a reference, and then compare the colors of the captured Moiré code with the reference image to decrypt the original QR code. For example, if the color difference of a given pixel in two images exceeded a given threshold (80, 120, 120 in the RGB channels, respectively), then the pixel is marked as a different color. If more than 80% of the pixels in a block are different, then the block is colored black; otherwise, the block is white (Fig. 17(c)).

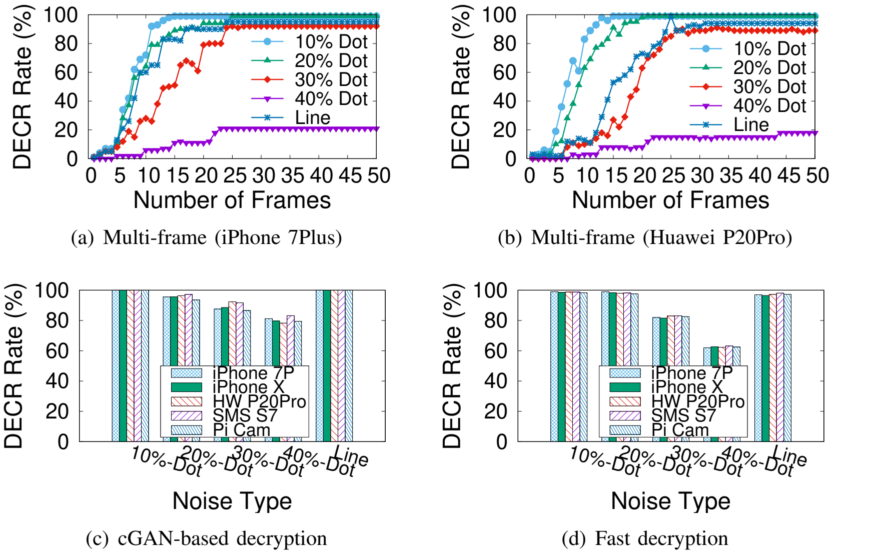
This article has been accepted for publication in IEEE Transactions on Dependable and Secure Computing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TDSC.2025.3598300



(a) Multi-frame-based

(b) cGAN-based

Fig. 18. Decryption rate vs. the number of frames



(a) Multi-frame (iPhone 7Plus)

(b) Multi-frame (Huawei P20Pro)

(c) cGAN-based decryption

(d) Fast decryption

Fig. 19. Decryption rates of MoiréComm with 10%-40% added dot-type noises or line-type noise

V. Evaluation

A. Experiment Setup and Dataset Collection

We generate 20 version-3 (29 × 29) QR codes to encode random text messages with the error correction level set at "M" (*i.e.*, 15% data restoration) [2] using MoiréComm for encryption. The generated Moiré codes are displayed on 10 displays (4 iOS, 4 Android, 1 desktop display, and 1 laptop display). Furthermore, the codes are configured specifically for receivers position at a specific distance at an angle of 0°. 8 smartphones (4 iOS and 4 Android) with built-in cameras and 2 PiCameras connected to a Raspberry Pi [33] are used to decrypt the Moiré codes. The cameras are set to record a video of each Moiré code at $30fps$ for $5sec$. Each experiment is repeated 30 times for each of the 20 Moiré codes. For the multi-frame decryption method, the screen is fixed and the camera is held by an user. For the fast decryption method, both the display and the camera are fixed. We report the averaged percentage of messages that are correctly extracted from the Moiré codes.

The proposed cGAN-based decryption model is a data-intensive model, we generate 100 QR code images for training and use the aforementioned 20 QR code images, that are unseen for the trained model, for testing. Considering the robustness of the cGAN-based decryption model on various display-camera pairs and designated positions, we captured the Moiré code images for the 100 QR code images with the screens fixed and the cameras held by the users. For each QR code image, we record the Moiré codes from each angle in [-5°, 5°] with one degree interval at fixed optimal distances and each distance in [-5 cm , 5 cm] with one centimeter interval at fixed optimal angles on the six cameras and six screens. The original QR code images and Moiré codes captured by the cameras are all resized as 256 × 256 pixels. Ultimately, we obtain dataset with **125000** pairs of (Moiré codes, QR code) for training the proposed cGAN-based decryption model, and 4800 pairs for testing the model performance on unseen Moiré codes and 4× 4 unseen screen-camera pairs. The entire process is trained on one Tesla V100 GPU, and takes approximately 37 hours with 3 GB video memory until the model converged.

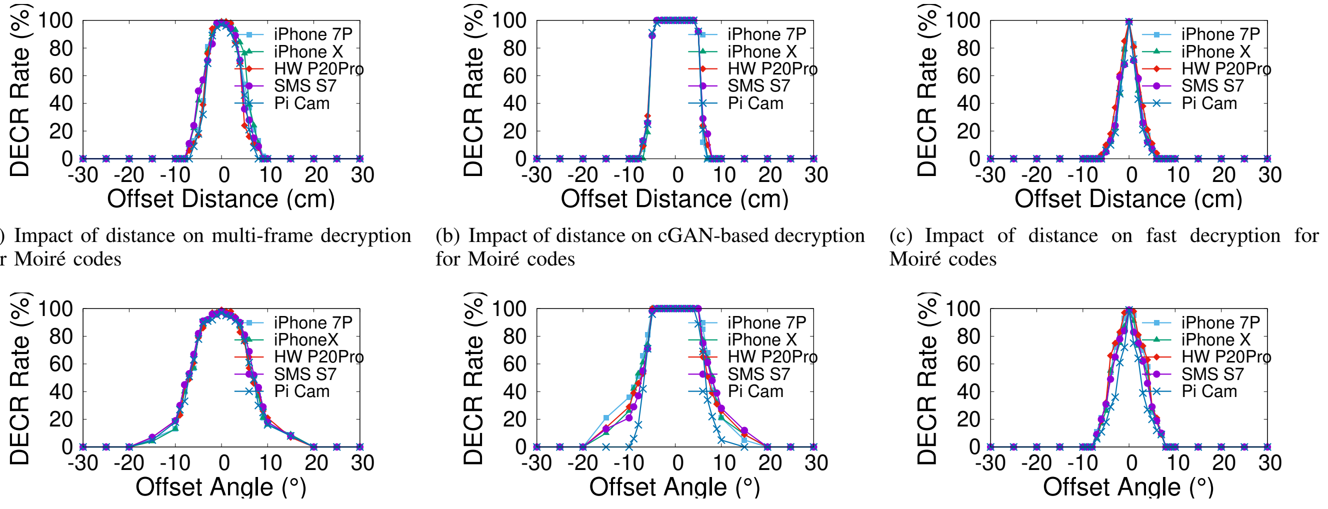
B. Micro Benchmark

1) Frames Required for Decryption: The quantity of frames needed for successful decryption is a critical metric in evaluating the practicality of the MoiréComm system. The expedited decryption method necessitates a single reference image along with one Moiré code image for the decryption process. Consequently, the ensuing assessment experiments are focused on scenarios involving handheld photography: specifically, the traditional multi-frame decryption approach and the cGAN-based multi-frame decryption methodology.

To determine the number of frames in the multi-frame decryption scheme, we first generate encrypted QR code images with the additional noise type as 10% dot (see Sec. **??**), then we use 5 cameras to record videos of 20 Moiré code at $30fps$ for 5 seconds and repeat 30 times. We apply the multi-frame decryption method to all videos and filter out the blurred frames. Fig. 18(a) shows the cumulative distribution function (CDF) of the number of frames required to decrypt Moiré codes when the camera is held in the correct position. We can see that the average number of frames required for multi-frame decryption is 10.2 and with 16 frames all Moiré codes can be correctly decrypted. In the following evaluation, we use 16 frames to evaluate multi-frame decryption performance if there are no special instructions.

Although our proposed cGAN-based decryption scheme is designed to input one captured Moiré code image to output the original QR code image, the experimental results show that our method also cannot deal with the blur parts in the Moiré code. Thanks to the error-correcting mechanism in the QR code generator, the embedded information can still be recovered from the output of the cGAN-based method. We evaluate the number of frames required by the cGAN-based method by

[2] Alipay uses version-2 (25×25) QR codes [32] while WeChat uses version-1 (21 × 21) [32] QR codes. Version-3 QR codes which carry more data are representative of the amount of data needed by these systems.



(a) Impact of distance on multi-frame decryption for Moiré codes

(b) Impact of distance on cGAN-based decryption for Moiré codes

(c) Impact of distance on fast decryption for Moiré codes

(d) Impact of angle on multi-frame decryption for Moiré codes

(e) Impact of angle on cGAN-based decryption for Moiré codes

(f) Impact of angle on fast decryption for Moiré codes

Fig. 20. Decryption rate of *Moiré* QR codes, where the screen is set as a coordinate system with its center as the origin

successively identifying the cGAN's method's output from the same photographing position until the original QR code embedded information is successfully scanned. Fig. 18(b) shows that the embedded information can be successfully recovered with nearly 100% probability after scanning three outputting results from the cGAN-based method. In the next experiment, we will continuously scan 3 frames output by the cGAN-based method to evaluate its decryption performance.

2) Determining the Additional Noise Type: The additional noises are designed to camouflage the boundaries resulting from abrupt phase changes in the encrypted QR code images. However, the type of these noises also affects the decryption efficiency of Moiré codes because noises indeed increase the error bit rate in the original QR codes.

We evaluate the impact of added dot-type noises or line-type noise on the three decryption schemes, and determine the best noise type for practical scenarios. Figs. 19(a)-19(b) show the decryption rates and corresponding number of frames required for the multi-frame decryption scheme after adding 10%-40% dot-type and line-type noises respectively. The decryption rate is above 95% when using average 15 frames for 10% dot noise and 23 frames for 20% dot noise. Nonetheless, the decryption rate do not improve with the addition of more frames. When 40% dot noise is added, decryption rates drop to less than 20%, even when using 50 frames. Investigating the traces that fails decryption reveals that the increasing dot noise and camouflaging lines add the errors in block color classification, which decrease the decryption rate. Fig. 19(d) and 19(c) show the impact of additional noise types on the average decryption rate when using the fast decryption scheme and cGAN-based decryption scheme, respectively. When only 10% or 20% dot noise is added, the decryption rates are almost unaffected and above 97%. When 40% added, the decryption rate drops to 61% for fast method and 78% for cGAN-based method.

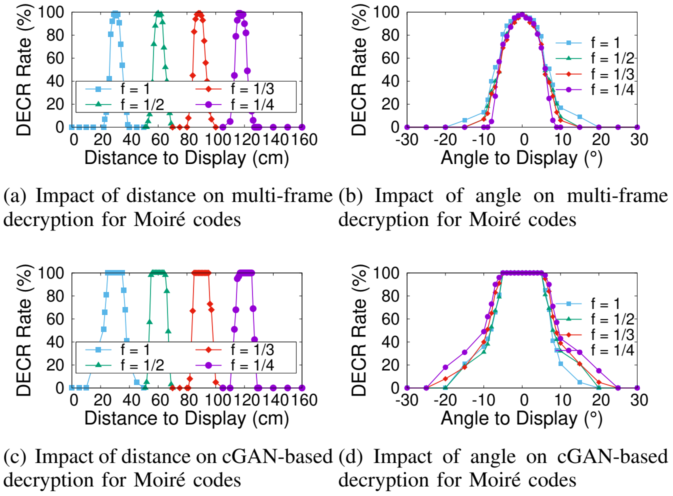
From the above experimental results, 10% dot-type added noise is sufficient to prevent the camera from capturing original QR code image's boundaries from the Moiré codes, while all our proposed decryption schemes can achieve satisfactory performance. In the following experiment, we adopt the additional 10% dot-type noise in the encryption process of the QR code images.

C. MoiréComm Performance

1) Decryption Range: The primary objective behind the development of MoiréComm is to enhance the security of camera-screen communication. Prior work [34] have showed the decryption rates of the standard QR codes can be achieved up to 100% while being placed within 3 meters from the display, and the view angle has little impact to the QR code decryption rate (100% ranging from -89° to 89°). These results imply that the standard QR codes are easily sniffed.

We then evaluate the decryption range of the aforementioned three decryption schemes in Sec. IV-C when decoding the captured Moiré codes. Here, we define the screen as a coordinate system with its center as the origin. The distance from the camera to this origin is set as the reception distance, and the angle of the lens facing the center is defined as the reception angle. To better understand the high security of our system in physical space, we use the offset distance and angle towards the reception area to measure the decryption rate of recipients for Moiré codes. Specifically, we define a legitimate receiving area, which is a uniquely specified point. Then, the offset distance is defined as the distance from the receiving camera to this point, and the offset angle is the deviation between the angle at which the receiving camera faces the screen and the standard angle (*i.e.*, 0°). Figs. 20(a), 20(b) and 20(c) present the decryption rates of Moiré codes with 5 cameras positioned at the correct view angle but at various distances from the screen. Figs. 20(d), 20(e) and 20(f) present the decryption rates with the camera positioned at the correct distance but at various view angles. When the camera is positioned at the designated distance (shifted by 0 cm) and at the designated angle (shifted by 0°), the decryption rate is 100%. When the camera is 10 cm or 20°away from the designated position, the decryption rate drops to 0. These results demonstrate the efficacy of MoiréComm in preventing

This article has been accepted for publication in IEEE Transactions on Dependable and Secure Computing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TDSC.2025.3598300



(a) Impact of distance on multi-frame decryption for Moiré codes

(b) Impact of angle on multi-frame decryption for Moiré codes

(c) Impact of distance on cGAN-based decryption for Moiré codes

(d) Impact of angle on cGAN-based decryption for Moiré codes

Fig. 21. Multi-frame and cGAN-based decryption rates with different frequency modulations

QR codes from being sniffed outside the designated receiver position areas.

Our research underscores that our proposed cGAN-based decryption technique achieves decryption success rates approaching 100% within a photographing range of ± 5 cm and ± 5° deviation from the target position. This high level of accuracy is attributed to the supervised nature of the cGAN-based method, with its training data encompassing Moiré codes captured within the specified reasonable photographing range. Nevertheless, the expansion of this range in the training data does not substantially enhance the decryption capability of the cGAN-based algorithm. This is because Moiré patterns recorded outside the ± 5 cm and ± 5° range lose clarity and progressively take on a “gray” appearance, as illustrated in Fig. 2(d), containing almost zero useful information about the original QR code.

2) Impact of Frequency Modulation: To evaluate the frequency modulation in the MoiréComm encryption process, we modulate 20 QR codes with four frequencies and display them on a DELL S2340M screen. We let the user hold an iPhone 7Plus and capture the Moiré codes at four designated positions, then we apply the multi-frame and cGAN-based decryption methods to evaluate the decryption performance. The corresponding results are shown in Fig. 21. First, we can see that when a smaller modulation frequency is used, the Moiré codes can be decrypted at a longer distance. It provides the flexibility in designing Moiré codes for applications targeted at various operating distances. Second, the Moiré codes still offer the high security since they can only be decrypted at the designated distances and angles. Third, our trained cGAN-based decryption model (encrypted images in the training datasets are encoded with $f = 1$) has similar decryption performances on the Moiré codes modulated at other frequencies.

3) Impact of Ambient Lights: To assess the impact of ambient light intensity on our system’s performance, we conducted two sets of experiments, one indoors and the other outdoors. The first set of experiments was carried out in an indoor setting, specifically in a nighttime environment, where a desk lamp with a power rating of 15 watts and a maximum brightness of approximately 1500 lumens was used. We adjusted the brightness levels of the lamp (*e.g.*, 0%, 20%, 50%, ..., 100%) to simulate various ambient light conditions. The second set of experiments took place outdoors during a sunny summer midday, chosen to test the scenario of high-intensity sunlight. Throughout the experiments, the QR code display screen (DELL S2340M) maintained maximum brightness, and the camera (iPhone 7 Plus) was set to an automatic adjustment mode. We tested 25 Moiré codes using three different decoding methods, capturing each code 20 times. The decoding rate results are presented in Fig. 22. The results indicate that the intensity of external ambient light has a minimal effect on the system’s performance, consistent with traditional QR code usage experience, where decoding success is primarily dependent on the brightness of the screen itself.

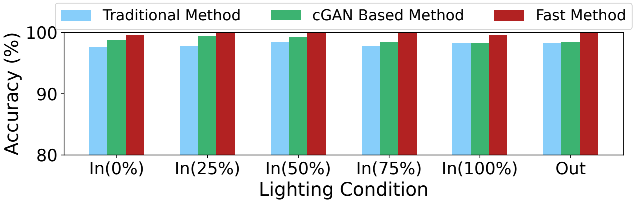
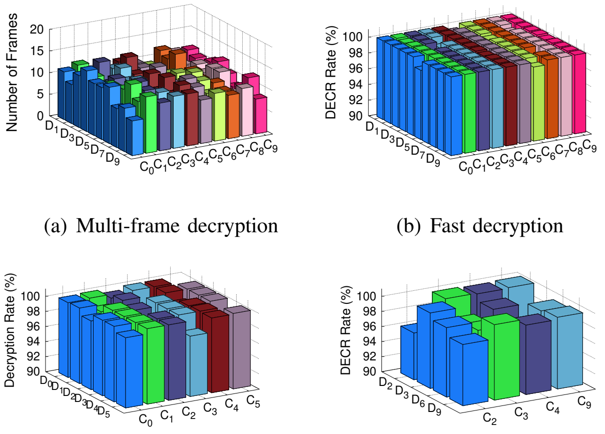


Fig. 22. Decryption performances of Moiré codes under different lighting conditions: ’In’ indicates indoors, with the percentage referring to the intensity at which a desk lamp, placed beside the QR code screen, is controlled for brightness. ’Out’ indicates outdoors, with sunny weather at noon

4) Impact of Screens and Cameras: We also examine how MoiréComm works on a variety of mobile devices. Ten devices are used to display Moiré codes and ten mobile cameras are used to capture videos to decrypt the Moiré codes. Fig. 23(a) shows the average number of frames required for multi-frame decryption. All of the display-camera pairs work in a similar manner, wherein an average of 11.3 frames is required for decryption. The average decryption rate under the fast decryption scheme is presented in Fig. 23(b). In these tests, MoiréComm is proved to be highly robust, with an average decryption rate of 98.6%. We evaluate the robustness of the cGAN-based decryption model proposed in Sec. IV-C3 with 20 unseen encrypted QR codes. On the trained screens and cameras, as shown in Fig. 23(c), our model can achieve an average decryption rate of 98.7% when capturing the Moiré codes around the designated positions; on the unseen screens and cameras, as shown in Fig. 23(d), an average decryption rate of 96.9% can be achieved. The above experiments also show that our cGAN-based decryption model provides good robustness when encountering unseen screens and cameras.

5) Consumption of the MoiréComm: In this part, we evaluate the efficiency of our proposed decryption schemes. Initially, we assess the optimal photographing range that enables users to decode Moiré codes with a decryption success rate of up to 95% using our three suggested decryption methods. A more extensive photographing range correlates with an improved user experience. According to the data presented in Table I, the cGAN-Based decryption scheme not only provides a considerable photographing area but also maintains



(a) Multi-frame decryption

(b) Fast decryption

(c) cGAN-based decryption on the trained devices

(d) cGAN-based decryption on the unseen devices

Fig. 23. Impact of various cameras and displays. Displays are D_0: iPhone 6; D_1: iPhone 7Plus; D_2: iPhone X; D_3: iPhone XS; D_4: Huawei P20Pro; D_5: Samsung S7; D_6: Nexus 6P; D_7: Google Pixel 2; D_8: DELL S2340M; D_9: MacBookPro 2016; Cameras are C_0: iPhone 6; C_1: iPhone 7Plus; C_2: iPhone X; C_3: iPhone XS; C_4: Huawei P20Pro; C_5: Samsung S7; C_6: Nexus 6P; C_7: Google Pixel 2; C_8: Pi Camera (5MP); P_9: Pi Camera (8MP)

a high rate of successful decryption. Furthermore, we analyze the time it takes to decrypt using the three methods. The decryption process was performed offline on a laptop with an Intel i7-10875H CPU at 2.80 GHz, 32 GB of RAM, and an NVIDIA GeForce RTX 2060 with 6 GB of VRAM. Also as indicated in Table I, the cGAN-Based decryption method achieves a minimal decryption time of only 0.02 s, albeit at the expense of greater VRAM usage. Given that the VRAM requirement (approximately 224.2 MB) is within the capabilities of contemporary mobile devices, we propose that the negligible decryption time achieved with the cGAN-based approach significantly enhances the practicality of our MoiréComm.

TABLE I
COMPARISON OF DECRYPTION SCHEMES

	Multi-frame	cGAN-Based	Fast
Dist. range	$[-2cm, 2cm]$	$[-4cm, 4cm]$	$[-0.5cm, 0.5cm]$
Angle range	$[-4°, 4°]$	$[-6°, 6°]$	$[-1°, 1°]$
DECR rate	98.6%	98.8%	99.4%
# of frames	16	3	2
DECR delay	2.7 ± 0.07s	0.02 ± 0.067s	0.40 ± 0.01s
(V)RAM	27.4MB	224.2MB	4.6MB
Shoot case	Handheld	Handheld	Fixed

VI. DISCUSSION

The display methods of QR codes are not limited to electronic screens, such as LEDs and OLEDs, and encompass a variety of other approaches. For instance, QR codes can be printed on paper to create flyers and posters, or they can be fashioned into artworks through 3D printing or engraving. Although these display methods differ, they essentially involve the visual rendering of each pixel. In comparing printed and electronic screen QR codes, the pixel units of an electronic screen correspond to the smallest discernible units that a printer can produce, determined by the printer's DPI (*i.e.*, Dots Per Inch). With a high-precision laser printer, it is possible to meticulously print each pixel of an encrypted QR code image onto paper. Consequently, even when captured by a camera, the encrypted QR code continues to convey the designed spatial frequency information and exhibits optical non-linear characteristics to generate the Moiré pattern. Therefore, our system is applicable to scenarios involving QR codes printed on paper. Similarly, for 3D printing or cutting, as long as pixel-level precision is achieved, our system remains applicable.

However, our system is primarily designed for QR code communication scenarios that require high levels of security, such as mobile payments, electronic social business cards, and access control systems. These QR codes need to have properties that allow dynamic updating, whereas QR codes printed on paper and those produced through 3D printing lack the capability for modification and updating once created. Therefore, for high-security QR code communication scenarios, we recommend users employ screen-displayed QR codes, utilizing our system to enhance the security of screen-camera communication.

VII. RELATED WORK

Near-field communications [35], such as, electromagnetic side-channel [36], [37], [38], [39], [40], [41], radio-frequency channel [36], [42], [43], and acoustic channel [44], [45], [46], [47], [48], [49], and screen-camera communication methods [50], [51], [52], is a widely utilized for the dissemination of information across various practical settings. Among them, screen-camera communication (*e.g.*, QR code) is the most popular one, that can be easily deployed on the screens and effortlessly captured using smartphone cameras. Despite its convenience, QR code is known to have security vulnerabilities as discussed in several studies [3], [4], [5], [6]. To address these concerns, researchers have developed several methods of visual and optical cryptography, which conceal data within disguised visual motifs. In this context, we provide an overview of the literature pertinent to the secure transmission of data using visual patterns. The prior research has primarily concentrated on three key areas: 1) the development of QR codes that facilitate efficient data encoding and decoding processes, 2) the implementation of visual and optical cryptographic techniques for secure data transmission, and 3) the application of Moiré patterns as a covert medium for message concealment.

A. Design of QR Codes

Conventional QR codes are composed of a series of black and white squares that encode specific data elements. Businesses and academic researchers have modified these codes by introducing slight variations in their patterns to include colors, logos, and additional elements, thus personalizing the codes for various applications, as noted in the works of [53] and [54]. Innovations such as ARTCode [55], halftone QR codes [56], and PiCode [57] have gone a step further by embedding information into content that is directly readable by humans, enhancing the overall user experience by making it more engaging and informative. To leverage the varied capabilities of contemporary hardware, Strata [58] is proposed as a multi-layered encoding approach that caters to various image capture resolutions, allowing for information to be transmitted at different levels of detail. Meanwhile, ScreenID [59], [60] leverages the pulse-width modulation (PWM) frequency of digital displays as a unique identifier to fortify the security of QR codes. The MoiréComm technology presents a complementary approach to these developments. That is, it can be seamlessly integrated with all of these schemes. Indeed, it is entirely possible to create a MoiréComm code that incorporates information in a format that is accessible and legible to human users.

B. Optical and Visual Cryptography

While the aforementioned studies have predominantly focused on enhancing the readability and encoding efficiency of QR codes, they have largely overlooked security concerns associated with QR code usage. Current applications of QR codes that demand secure communication, such as mobile payments [61] and authentication mechanisms [62], [63], [64], typically incorporate encryption directly within the QR codes themselves. However, recent research indicates that mere encryption of messages within QR codes is insufficient in countering security threats. This is because attackers executing Replay and Secure Transaction Lock Screen attacks require only the QR code image and do not need to decrypt the message content to carry out their exploits [6].

To enable the complete concealment of visual images, existing optical visual cryptography (O VC) techniques [65] encode a secret image into shared images with camouflaged visual patterns such that stacking a sufficient number of shared images reveals the original secret image. In [66], [67], [62], [68], [69], VC technology has been applied to QR codes to check the identity of individuals accessing QR codes or to control permissions related to accessing protected data. However, those works require that users scan multiple images or exchange key images in advance to recover the original QR code. In contrast, the scheme proposed in this study requires only that users hold the camera in a designated position to immediately obtain embedded messages.

Optical encryption is another viable technique for concealing images. Double Random Phase Encoding method [70], along with its numerical counterparts [71], [72], [73], [74], [75], employs a sequence of optical components, such as lenses, to encrypt images. Utilization of these optical technologies for the concealment of encrypted QR codes has been explored in studies [76], [77], [78]. However, these optical encryption techniques necessitate the use of specialized optical equipment, in contrast to the mQR code system, which decrypts hidden QR codes using the ubiquitous camera found in standard smartphones.

C. Image Steganography

Image steganography is a method that hides information in images, aiming to ensure embedded secrets are imperceptible. Unlike OVC, which focuses on encryption and recovery, steganography requires images to look completely normal without visible anomalies. Traditional image steganography primarily includes three types of techniques: spatial-based methods [79], [80], transform-based methods [81], [82], segmentation-based methods [83], [84], and adaptive steganography methods [85], [86]. In recent years, various deep learning-based image steganography schemes have been introduced. These methods can be categorized into four types [87]: synthesis-based [88], probability map generation-based [89], adversarial embedding-based [90], and 3-player game-based [91], [92].

These image steganography techniques are primarily applied in copyright protection [93], digital image communication [94], integrity verification [95], and so on. The common goal in these applications is to ensure that recipients can fully receive image information processed through steganography, meaning these images are transmitted in a lossless manner (such as wired or wireless communication with lossless protocols), allowing legitimate recipients to decode the hidden content.

In the secure QR code communication scenario, image information is transmitted through the optical channel between the screen and the camera. This means the information captured by the camera from the screen is actually lossy, and the camera cannot capture the original information of each pixel displayed in an image encrypted through steganography. This is related to the color gamut of the screen as well as the camera's lens CFA. Therefore, image steganography methods cannot be applied to QR code communication scenarios to enhance security.

D. Moiré-based Technologies

Moiré patterns have been utilized in a variety of studies to conceal imagery. The work of [96] investigates the generation of Moiré patterns through the overlay of grating patterns, specifically to produce facial images that can be perceived by the human eye. [97] demonstrates the creation of dynamic Moiré effects, which manifest as moving elements that vary in speed and direction upon the translation of an overlay layer. [98] introduces a system for the covert sharing of information within realistic images. Additionally, [20] presents a technique to craft Moiré art that can be visually decoded by superimposing printed grating images on separate transparencies. Common to these techniques is the requirement of two semi-transparent layers to be superimposed in order to disclose the concealed image. Beyond image concealment, novel applications of Moiré patterns have been proposed, such as an innovative snapshot approach for angular measurement and tracking [99], and a straightforward yet effective framework [100] for the extraction of Moiré edge maps from images affected by Moiré patterns. Additionally, [101] proposes a watermark-like method capable of generating a deliberate Moiré pattern on a photograph when captured towards a screen.

Our work is inspired [20], but, unlike this study, we design phase modulation and frequency modulation methods to achieve more covert encryption of QR codes and support various secure distance communication scenarios. Additionally, we propose a series of algorithms for decoding the captured Moiré QR codes, considering the different computational capabilities and application scenarios of end-side devices. Compared to our previous work[102], [103], [34], in this paper, we propose a decoding method based on cGAN, which achieves a decoding delay of 0.02 seconds and an accuracy rate of up to 98% on end-side camera devices with certain computational resources. This significantly enhances the performance and efficiency of the system, providing a user experience similar to traditional QR codes.

VIII. Conclusion

We present MoiréComm, a system designed for highly secure and dependable camera-screen communication. MoiréComm offers several advantages over existing technologies. Firstly, it is a software-centric approach, eliminating the need for additional hardware or alternative communication methods. Secondly, the Moiré-based encryption and decryption processes are contingent upon the spatial relationship between the Moiré code and the capturing camera, creating a physical barrier for attackers who cannot access the same location as the intended victim. Lastly, with a decryption time of up to 0.02 s for the captured Moiré codes, users can enjoy a seamless scanning and decoding experience that is comparable to the traditional QR codes in real-world scenarios.

References

[1] Alipay, “Alipay: Experience fast, easy and safe online payments,” 2019. [Online]. Available: https://intl.alipay.com/

[2] AliPay, “Alipay service (merchant scans with integrator and acquirer),” 2019. [Online]. Available: https://global.alipay.com/service/barcode/7

[3] S. Sung, J. Lee, J. Kim, J. Mun, and D. Won, “Security analysis of mobile authentication using qr-codes,” vol. 5, 2015.

[4] H. Keni, M. Earle, and M. Min, “Product authentication using hash chains and printed qr codes,” in *2017 14th IEEE Annual Consumer Communications & Networking Conference (CCNC)*. IEEE, 2017, pp. 319–324.

[5] V. Mavroeidis and M. Nicho, “Quick response code secure: A cryptographically secure anti-phishing tool for qr code attacks,” in *International Conference on Mathematical Methods, Models, and Architectures for Computer Network Security*. Springer, 2017, pp. 313–324.

[6] X. Bai, Z. Zhou, X. Wang, Z. Li, X. Mi, N. Zhang, T. Li, S.-M. Hu, and K. Zhang, “Picking up my tab: Understanding and mitigating synchronized token lifting and spending in mobile payment,” in *26th USENIX Security Symposium (USENIX Security 17)*, 2017, pp. 593–608.

[7] B. O’Donnell, “Steals money sneakily by scanning people’s qr code.” 2019. [Online]. Available: https://www.thatsmags.com/shanghai/post/27482/man-steals-money-sneakily-scanning-people-s-qr-codes-in-shanghai

[8] T. Li, “Qr code scams rise in china, putting e-payment security in spotlight.” 2019. [Online]. Available: https://www.jianshu.com/p/b9657161933a

[9] P. Shankdhar, “Security attacks via malicious qr codes,” 2015. [Online]. Available: https://resources.infosecinstitute.com/security-attacks-via-malicious-qr-codes/

[10] T. Vidas, E. Owusu, S. Wang, C. Zeng, L. F. Cranor, and N. Christin, “Qrishing: The susceptibility of smartphone users to qr code phishing attacks,” in *International Conference on Financial Cryptography and Data Security*. Springer, 2013, pp. 52–69.

[11] D. P. Mitchell and A. N. Netravali, “Reconstruction filters in computer-graphics,” in *ACM Siggraph Computer Graphics*, vol. 22, no. 4. ACM, 1988, pp. 221–228.

[12] C. S. Williams and O. A. Becklund, *Introduction to the optical transfer function*. Wiley New York etc, 1989.

[13] Lensora, “Lenses with the longest focal length,” 2019. [Online]. Available: http://www.lensora.com/list_lenses.asp?sel=zoom_max

[14] Wikipedia, “List of longest smartphone telephoto lenses,” 2025. [Online]. Available: https://en.wikipedia.org/wiki/List_of_longest_smartphone_telephoto_lenses

[15] D. Industry, “Time-of-flight image sensor imx858,” 2025. [Online]. Available: https://www.directindustry.com/prod/sony-semiconductors/product-29226-2707153.html

[16] DPReview, “Nikon d7000 specs,” 2025. [Online]. Available: https://www.dpreview.com/products/nikon/slrs/nikon_d7000/specifications

[17] E. W. Weisstein, “Convolution,” 2003.

[18] I. Amidror, *The Theory of the Moiré Phenomenon: Volume I: Periodic Layers*. Springer Science & Business Media, 2009, vol. 38.

[19] Wikipedia, “Wikipedia, color filter array,” 2019. [Online]. Available: https://en.wikipedia.org/wiki/Bayer_filter/

[20] P.-H. Tsai and Y.-Y. Chuang, “Target-driven moire pattern synthesis by phase modulation,” in *Proceedings of the IEEE International Conference on Computer Vision*, 2013, pp. 1912–1919.

[21] A. Raney, “Pinhole camera theory summary,” 2017. [Online]. Available: https://ourpastimes.com/pinhole-camera-theory-summary-12210465.html

[22] A. Davies and P. Fennessy, *Digital imaging for photographers*. Focal Press, 2012.

[23] S. Farsiu, M. Elad, and P. Milanfar, “Multi-frame demosaicing and super-resolution of color images,” California Univ Santa Cruz Electrical Engineering Dept, Tech. Rep., 2006.

[24] K. Hirakawa and P. J. Wolfe, “Spatio-spectral color filter array design for optimal image recovery,” *IEEE Transactions on Image Processing*, vol. 17, no. 10, pp. 1876–1890, 2008.

[25] Y. Huang and Y. Long, “Demosaicking recognition with applications in digital photo authentication based on a quadratic pixel correlation model.” IEEE, 2008, pp. 1–8.

[26] K. J. Connolly, *The psychobiology of the hand*. Cambridge University Press, 1998.

[27] PyTorch, “Pytorch mobile,” 2023. [Online]. Available: https://pytorch.org/mobile/android/

[28] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, “Generative adversarial nets,” *Advances in neural information processing systems*, vol. 27, 2014.

[29] Y. Fu, S. Wang, L. Zhong, L. Chen, J. Ren, and Y. Zhang, “Svoice: Enabling voice communication in silence via acoustic sensing on commodity devices,” in *Proceedings of the 20th ACM Conference on Embedded Networked Sensor Systems*, 2022, pp. 622–636.

[30] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, “Image-to-image translation with conditional adversarial networks,” *CVPR*, 2017.

[31] Barcodes, “Qr code barcode scanner,” 2019. [Online]. Available: https://www.barcodesinc.com/cats/barcode-scanners/qr.htm

[32] J. Steeman, “Qr code data capacity,” 2019. [Online]. Available: https://blog.qr4.nl/page/QR-Code-Data-Capacity.aspx

[33] R. P. Foundation, “Raspberry pi,” 2019. [Online]. Available: https://www.raspberrypi.org/

[34] H. Pan, Y.-C. Chen, L. Yang, G. Xue, C.-W. You, and X. Ji, “mqrcode: Secure qr code using nonlinearity of spatial frequency in light,” in *The 25th Annual International Conference on Mobile Computing and Networking*, 2019, pp. 1–18.

[35] V. Coskun, K. Ok, and B. Ozdenizci, *Near field communication (NFC): From theory to practice*. John Wiley & Sons, 2011.

[36] N. H. Motlagh, “Near field communication (nfc)-a technical overview,” in *University of VAASA*, 2012.

[37] H. Pan, Y.-C. Chen, G. Xue, and X. Ji, “Magnecomm: Magnetometer-based near-field communication,” in *Proceedings of the 23rd Annual International Conference on Mobile Computing and Networking*, ser. MobiCom ’17. ACM, 2017, pp. 167–179.

[38] G. Xue, H. Pan, Y.-C. Chen, X. Ji, and J. Yu, “Magnecomm+: Near-field electromagnetic induction communication with magnetometer,” *IEEE Transactions on Mobile Computing*, vol. 22, no. 5, pp. 2789–2801, 2021.

[39] H. Pan, F. Tan, W. Li, Y.-C. Chen, L. Yang, G. Xue, and X. Ji, “Magdefender: Detecting eavesdropping on mobile devices using the built-in magnetometer,” in *2022 19th Annual IEEE International Conference on Sensing, Communication, and Networking (SECON)*. IEEE, 2022, pp. 28–36.

[40] H. Pan, L. Yang, H. Li, C.-W. You, X. Ji, Y.-C. Chen, Z. Hu, and G. Xue, “Magthief: Stealing private app usage data on mobile devices via built-in magnetometer,” in *2021 18th Annual IEEE International Conference on Sensing, Communication, and Networking (SECON)*. IEEE, 2021, pp. 1–9.

[41] Y. Fu, L. Yang, H. Pan, Y.-C. Chen, G. Xue, and J. Ren, "Magspy: Revealing user privacy leakage via magnetometer on mobile devices," *IEEE Transactions on Mobile Computing*, 2024.

[42] K. Finkenzeller, *RFID handbook: fundamentals and applications in contactless smart cards, radio frequency identification and near-field communication*. John wiley & sons, 2010.

[43] W. Choi, J.-S. Kim, J.-H. Bae, G. Choi, and J.-S. Chae, "Near-field antenna for a radio frequency identification shelf in the uhf band," *IET microwaves, antennas & propagation*, vol. 4, no. 10, pp. 1538–1542, 2010.

[44] H. Pan, Y.-C. Chen, Q. Ye, and G. Xue, "Magicinput: Training-free multi-lingual finger input system using data augmentation based on mnists," in *Proceedings of the 20th International Conference on Information Processing in Sensor Networks (co-located with CPS-IoT Week 2021)*, 2021, pp. 119–131.

[45] H. Pan, Y. Fu, Y. Qi, Y.-C. Chen, and J. Ren, "Magicwrite: One-dimensional acoustic tracking-based air writing system," *IEEE Transactions on Mobile Computing*, 2025.

[46] Y. Zhang, H. Pan, D. Ding, Y. Pan, Y.-C. Chen, L. Qiu, G. Xue, T. Chen, and X. Zhang, "Swifttrack+: Fine-grained and robust fast hand motion tracking using acoustic signal," *IEEE/ACM Transactions on Networking*, 2024.

[47] Y. Zhang, H. Pan, Y.-C. Chen, L. Qiu, Y. Lu, G. Xue, J. Yu, F. Lyu, and H. Wang, "Addressing practical challenges in acoustic sensing to enable fast motion tracking," in *Proceedings of the 22nd International Conference on Information Processing in Sensor Networks*, 2023, pp. 82–95.

[48] Y. Fu, Y. Zhang, Y. Lu, L. Qiu, Y.-C. Chen, Y. Wang, M. Wang, Y. Li, J. Ren, and Y. Zhang, "Adaptive metasurface-based acoustic imaging using joint optimization," in *Proceedings of the 22nd Annual International Conference on Mobile Systems, Applications and Services*, 2024, pp. 492–504.

[49] Y. Fu, Y. Zhang, H. Pan, Y. Lu, X. Li, L. Chen, J. Ren, X. Li, X. Zhang, and Y. Zhang, "Pushing the limits of acoustic spatial perception via incident angle encoding," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 8, no. 2, pp. 1–28, 2024.

[50] J. Zhao and X.-Y. Li, "Scsec: A secure near field communication system via screen camera communication," *IEEE Transactions on Mobile Computing*, vol. 19, no. 8, pp. 1943–1955, 2019.

[51] W. Hu, H. Gu, and Q. Pu, "Lightsync: Unsynchronized visual communication over screen-camera links," in *Proceedings of the 19th annual international conference on Mobile computing & networking*, 2013, pp. 15–26.

[52] K. Qian, Y. Lu, Z. Yang, K. Zhang, K. Huang, X. Cai, C. Wu, and Y. Liu, "{AIRCODE}: Hidden {Screen-Camera} communication on an invisible and inaudible dual channel," in *18th USENIX Symposium on Networked Systems Design and Implementation (NSDI 21)*, 2021, pp. 457–470.

[53] Roger, "Marc jacobs qr code," 2009. [Online]. Available: https://2d-code.co.uk/marc-jacobs-qr-code/

[54] R. Cox, "Qart codes," 2012. [Online]. Available: https://research.swtch.com/qr/draw

[55] Z. Yang, Y. Bao, C. Luo, X. Zhao, S. Zhu, C. Peng, Y. Liu, and X. Wang, "Artcode: preserve art and code in any image," in *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 2016, pp. 904–915.

[56] H.-K. Chu, C.-S. Chang, R.-R. Lee, and N. J. Mitra, "Halftone qr codes," *ACM Transactions on Graphics (TOG)*, vol. 32, no. 6, p. 217, 2013.

[57] W. Huang and W. H. Mow, "Picode: 2d barcode with embedded picture and vicode: 3d barcode with embedded video," in *Proceedings of the 19th annual international conference on Mobile computing & networking*. ACM, 2013, pp. 139–142.

[58] W. Hu, J. Mao, Z. Huang, Y. Xue, J. She, K. Bian, and G. Shen, "Strata: layered coding for scalable visual communication," in *Proceedings of the 20th annual international conference on Mobile computing and networking*. ACM, 2014, pp. 79–90.

[59] Y. Li, Y.-C. Chen, X. Ji, H. Pan, L. Yang, G. Xue, and J. Yu, "Screenid: Enhancing qrcode security by fingerprinting screens," in *IEEE INFOCOM 2021-IEEE Conference on Computer Communications*. IEEE, 2021, pp. 1–10.

[60] ——, "Toward a secure qr code system by fingerprinting screens," in *Proceedings of the 26th Annual International Conference on Mobile Computing and Networking*, 2020, pp. 1–3.

[61] S. Nseir, N. Hirzallah, and M. Aqel, "A secure mobile payment system using qr code," in *2013 5th International Conference on Computer Science and Information Technology*. IEEE, 2013, pp. 111–114.

[62] J. Lu, Z. Yang, L. Li, W. Yuan, L. Li, and C.-C. Chang, "Multiple schemes for mobile payment authentication using qr code and visual cryptography," *Mobile Information Systems*, vol. 2017, 2017.

[63] Y.-W. Chow, W. Susilo, G. Yang, M. H. Au, and C. Wang, "Authentication and transaction verification using qr codes with a mobile device," in *International Conference on Security, Privacy and Anonymity in Computation, Communication and Storage*. Springer, 2016, pp. 437–451.

[64] G. R. Corporation, "Secure quick reliable login," 2019, https://www.grc.com/sqrl/sqrl.htm.

[65] P. Punithavathi and S. Geetha, "Visual cryptography: A brief survey," *Information Security Journal: A Global Perspective*, vol. 26, no. 6, pp. 305–317, 2017.

[66] G. Horng, T. Chen, and D.-S. Tsai, "Cheating in visual cryptography," *Designs, Codes and Cryptography*, vol. 38, no. 2, pp. 219–236, 2006.

[67] S. Thamer and B. Ameen, "A new method for ciphering a message using qr code," *Comput. Sci. Eng.*, vol. 6, no. 2, pp. 19–24, 2016.

[68] W.-P. Fang, "Offline qr code authorization based on visual cryptography," in *2011 Seventh International Conference on Intelligent Information Hiding and Multimedia Signal Processing*. IEEE, 2011, pp. 89–92.

[69] X. Cao, L. Feng, P. Cao, and J. Hu, "Secure qr code scheme based on visual cryptography," in *2016 2nd International Conference on Artificial Intelligence and Industrial Engineering (AIIE 2016)*. Atlantis Press, 2016.

[70] G. Unnikrishnan, J. Joseph, and K. Singh, "Optical encryption by double-random phase encoding in the fractional fourier domain," *Optics letters*, vol. 25, no. 12, pp. 887–889, 2000.

[71] G. Situ and J. Zhang, "Multiple-image encryption by wavelength multiplexing," *Optics letters*, vol. 30, no. 11, pp. 1306–1308, 2005.

[72] A. Alfalou and C. Brosseau, "Optical image compression and encryption methods," *Advances in Optics and Photonics*, vol. 1, no. 3, pp. 589–636, 2009.

[73] M. Madjarova, M. Kakuta, M. Yamaguchi, and N. Ohyama, "Optical implementation of the stream cipher based on the irreversible cellular automata algorithm," *Optics letters*, vol. 22, no. 21, pp. 1624–1626, 1997.

[74] J. Han, C.-S. Park, D.-H. Ryu, and E.-S. Kim, "Optical image encryption based on xor operations," *Optical Engineering*, vol. 38, no. 1, pp. 47–55, 1999.

[75] E. Tajahuerce, O. Matoba, S. C. Verrall, and B. Javidi, "Optoelectronic information encryption with phase-shifting interferometry," *Applied Optics*, vol. 39, no. 14, pp. 2313–2320, 2000.

[76] J. F. Barrera, A. Mira, and R. Torroba, "Optical encryption and qr codes: secure and noise-free information retrieval," *Optics express*, vol. 21, no. 5, pp. 5373–5378, 2013.

[77] P. Cheremkhin, V. Krasnov, V. Rodin, and R. Starikov, "Qr code optical encryption using spatially incoherent illumination," *Laser Physics Letters*, vol. 14, no. 2, p. 026202, 2017.

[78] S. Jiao, W. Zou, and X. Li, "Qr code based noise-free optical encryption and decryption of a gray scale image," *Optics Communications*, vol. 387, pp. 235–240, 2017.

[79] D. N. Tran, H.-J. Zepernick, and T. M. C. Chu, "Lsb data hiding in digital media: a survey," *EAI Endorsed Transactions on Industrial Networks and Intelligent Systems*, pp. 1–50, 2022.

[80] P. Tsai, Y.-C. Hu, and H.-L. Yeh, "Reversible image hiding scheme using predictive coding and histogram shifting," *Signal processing*, vol. 89, no. 6, pp. 1129–1143, 2009.

[81] J. Fridrich, T. Pevnỳ, and J. Kodovskỳ, "Statistically undetectable jpeg steganography: dead ends challenges, and opportunities," in *Proceedings of the 9th workshop on Multimedia & security*, 2007, pp. 3–14.

[82] P. Sallee, "Model-based steganography," in *International workshop on digital watermarking*. Springer, 2003, pp. 154–167.

[83] K. Koptyra and M. R. Ogiela, "Steganography in qr codes—information hiding with suboptimal segmentation," *Electronics*, vol. 13, no. 13, p. 2658, 2024.

[84] ——, "Information hiding in qr codes using segment manipulation," in *2024 IEEE International Conference on Pervasive Computing and Communications Workshops and other Affiliated Events (PerCom Workshops)*. IEEE, 2024, pp. 397–400.

[85] L. Guo, J. Ni, and Y. Q. Shi, "An efficient jpeg steganographic scheme using uniform embedding," in *2012 IEEE International Workshop on Information Forensics and Security (WIFS)*. IEEE, 2012, pp. 169–174.

[86] T. Pevnỳ, T. Filler, and P. Bas, "Using high-dimensional image models to perform highly undetectable steganography," in *Information Hiding: 12th International Conference, IH 2010, Calgary, AB, Canada, June 28-30, 2010, Revised Selected Papers 12*. Springer, 2010, pp. 161–177.

[87] M. Chaumont, "Deep learning in steganography and steganalysis from 2015 to 2018," *arXiv preprint arXiv:1904.01444*, 2019.

[88] H. Shi, J. Dong, W. Wang, Y. Qian, and X. Zhang, "Ssgan: Secure steganography based on generative adversarial networks," in *Advances in Multimedia Information Processing–PCM 2017: 18th Pacific-Rim Conference on Multimedia, Harbin, China, September 28-29, 2017, Revised Selected Papers, Part I 18*. Springer, 2018, pp. 534–544.

[89] W. Tang, S. Tan, B. Li, and J. Huang, "Automatic steganographic distortion learning using a generative adversarial network," *IEEE Signal Processing Letters*, vol. 24, no. 10, pp. 1547–1551, 2017.

[90] W. Tang, B. Li, S. Tan, M. Barni, and J. Huang, "Cnn-based adversarial embedding for image steganography," *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 8, pp. 2074–2087, 2019.

[91] S. Baluja, "Hiding images within images," *IEEE transactions on pattern analysis and machine intelligence*, vol. 42, no. 7, pp. 1685–1697, 2019.

[92] J. Zhu, R. Kaplan, J. Johnson, and L. Fei-Fei, "Hidden: Hiding data with deep networks," in *Proceedings of the European conference on computer vision (ECCV)*, 2018, pp. 657–672.

[93] X. Zhang, R. Li, J. Yu, Y. Xu, W. Li, and J. Zhang, "Editguard: Versatile image watermarking for tamper localization and copyright protection," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2024, pp. 11 964–11 974.

[94] A. S. Ansari, "A review on the recent trends of image steganography for vanet applications." *Computers, Materials & Continua*, vol. 78, no. 3, 2024.

[95] P. Capasso, G. Cattaneo, and M. De Marsico, "A comprehensive survey on methods for image integrity," *ACM Transactions on Multimedia Computing, Communications and Applications*, vol. 20, no. 11, pp. 1–34, 2024.

[96] G. Lebanon and A. M. Bruckstein, "Variational approach to moiré pattern synthesis," *JOSA A*, vol. 18, no. 6, pp. 1371–1382, 2001.

[97] R. D. Hersch and S. Chosson, "Band moiré images," in *ACM Transactions on Graphics (TOG)*, vol. 23, no. 3. ACM, 2004, pp. 239–247.

[98] Y. Desmedt and T. Van Le, "Moiré cryptography," in *Proceedings of the 7th ACM conference on Computer and communications security*, 2000, pp. 116–124.

[99] S. Qiu, H. Amata, and W. Heidrich, "Moirétag: Angular measurement and tracking with a passive marker," in *ACM SIGGRAPH 2023 Conference Proceedings*, 2023, pp. 1–10.

[100] C. Yang, Z. Yang, Y. Ke, T. Chen, M. Grzegorzek, and J. See, "Doing more with moiré pattern detection in digital photos," *IEEE Transactions on Image Processing*, vol. 32, pp. 694–708, 2023.

[101] Y. Cheng, X. Ji, L. Wang, Q. Pang, Y.-C. Chen, and W. Xu, "{mID}: Tracing screen photos via {Moiré} patterns," in *30th USENIX Security Symposium (USENIX Security 21)*, 2021, pp. 2969–2986.

[102] H. Pan, Y.-C. Chen, G. Xue, C.-W. B. You, and X. Ji, "Secure qr code scheme using nonlinearity of spatial frequency," in *Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers*. ACM, 2018, pp. 207–210.

[103] H. Pan, Y.-C. Chen, and G. Xue, "Poster: Secure visible light communication via two-dimensional spatially aliased patterns," in *Proceedings of the 13th International Workshop on Wireless Network Testbeds, Experimental Evaluation & Characterization*, 2019, pp. 51–52.

Hao Pan is a Senior Researcher at Microsoft Research Asia (Shanghai). Before that, he received his Ph.D. in computer science at the Shanghai Jiao Tong University (SJTU) in 2022, and completed undergraduate studies in the Yingcai Honors College of University of Electronic Science and Technology of China (UESTC) in 2016. His research interests focus on networked systems and span the areas of wireless communication and sensing, human-comuputer interaction, and computer vision.

Yongjian Fu received the B.Sc. in Computer Science from Central South University in 2021, China. He is currently working toward a Ph.D. degree in Computer Science at the School of Computer Science and Engineering from Central South University. He is also a visiting PhD student since 2021 at the Department of Computer Science and Technology, Tsinghua University. His research interests include mobile computing, edge intelligence, as well as security and privacy.

Yu Lu is a PhD student in the School of Computer Science at Shanghai Jiao Tong University, where he has been studying since 2022. Prior to this, he received his B.Eng. degree from the School of Cyber Science and Engineering at Shanghai Jiao Tong University. His current research focuses on the Internet of Things (IoT), big data, mobile computing, and security.

Feitong Tan is a Research Scientist at Google XR, focusing on Digital Humans. I received my Ph.D. in 2021 from Simon Fraser University, where he was advised by Prof. Ping Tan. Earlier, he completed his undergraduate studies at the Yingcai Honors College, University of Electronic Science and Technology of China, in 2016. His research interests span computer vision and machine learning, with an emphasis on 3D reconstruction, generation, and human digitization.

Yi-Chao Chen joined Shanghai Jiao Tong University as a tenure-track Associate Professor in the Department of Computer Science and Engineering in 2018. He received his B.S. and M.S. in the Department of Computer Science and Information Engineering at National Taiwan University in 2004 and 2006, respectively. He got his Ph.D. in Computer Science at the University of Texas at Austin in 2015. Prior to joining SJTU, He spent a year as a Researcher in Huawei Future Network Theory Lab in Hong Kong and then worked as a Chief Scientist in Hauoli LLC. His research interests focus on networked systems and span the areas of mobile computing, wireless networking, and cyber-security.

Ju Ren (Senior Member, IEEE) received the BSc, MSc, and PhD degrees all in computer science, from Central South University, China. Currently, he is an associate professor with the Department of Computer Science and Technology, Tsinghua University, China. His research interests include Internet-of-Things, edge computing, edge intelligence, as well as security and privacy. He currently serves as an associate editor for many journals, including IEEE/ACM Transactions on Networking, IEEE Transactions on Mobile Computing, IEEE Transactions on Cloud Computing and IEEE Transactions on Vehicular Technology, etc. He also served as the general co-chair for IEEE BigDataSE'20, the TPC co-chair for IEEE BigDataSE'19, the track co-chair for IEEE ICDCS'24, the poster co-chair for IEEE MASS'18, a symposium co-chair for IEEE/CIC ICCC'23&19, I-SPAN'18 and IEEE VTC'17 Fall, etc. He received several best paper awards from IEEE flagship conferences, including IEEE ICC'19 and IEEE HPCC'19, etc., the IEEE TCSC Early Career Researcher Award (2019), and the IEEE ComSoc Asia-Pacific Best Young Researcher Award (2021). He was recognized as a highly cited researcher by Clarivate (2020-2022).