

OAcod: Overall Aesthetic 2D Barcode on Screen

Zehua Ma, Xi Yang, Han Fang, Weiming Zhang, and Nenghai Yu

Abstract—Nowadays, two-dimensional (2D) barcodes have been widely used in various domains. And a series of aesthetic 2D barcode schemes have been proposed to improve the visual quality and readability of 2D barcodes for better integration with marketing materials. Yet we believe that the existing aesthetic 2D barcode schemes are *partially* aesthetic because they only beautify the data area but retain the position detection patterns with the black-white appearance of traditional 2D barcode schemes. Thus, in this paper, we propose the first *overall* aesthetic 2D barcode scheme, called OAcod, in which the position detection pattern is canceled. Its detection process is based on the pre-designed symmetrical data area of OAcod, whose symmetry could be used as the calibration signal to restore the perspective transformation in the barcode scanning process. Moreover, an enhanced demodulation method is proposed to resist the lens distortion common in the camera-shooting process. The experimental results illustrate that when $5 \times 5\text{cm}$ OAcod is captured with a resolution of 720×1280 pixels, at the screen-camera distance of 10cm and the angle less or equal to 25° , OAcod has 100% detection rate and 99.5% demodulation accuracy. For $10 \times 10\text{cm}$ OAcod, it could be extracted by consumer-grade mobile phones at a distance of 90cm with around 90% accuracy.

Index Terms—aesthetic 2D barcode, perspective distortion, lens distortion, auto-convolution function

I. INTRODUCTION

WITH the increasing popularity of smart mobile devices and wireless network infrastructures [1], two-dimensional (2D) barcodes are used in a wide range of industries, from manufacturing and logistics to mobile marketing and business. Especially in mobile marketing, 2D barcodes play a significant role since 2008 [2], [3]. According to a research statistic from 2012 [4], print media, i.e., magazines, flyers, and newspapers, are the most frequently used media to present 2D barcodes, whose proportion is over 80%. Yet in 2022 [5], over 39% of 2D barcodes are scanned from screens, including TV and website on PC. People are increasingly preferring to scan 2D barcodes on screen because, on the one hand, the screen content grows explosively [6], and on the other hand, the restriction of physical contact during the COVID-19 pandemic and the implementation of the paperless office contribute to a rapid decline in the number of print media [7], [8]. Moreover, as media for mobile marketing, the screen has some advantages over print ones, such as higher color accuracy and better maintainability. By attaching 2D

Zehua Ma, Xi Yang, Weiming Zhang, and Nenghai Yu are with CAS Key Laboratory of Electromagnetic Space Information, University of Science and Technology of China, Hefei, 230026, China. Han Fang is with School of Computing, National University of Singapore, 117417, Singapore. Corresponding author: Weiming Zhang. (e-mail: mzh045@mail.ustc.edu.cn, zhangwm@ustc.edu.cn)

This work was supported in part by the Natural Science Foundation of China under Grant 62072421, 62102386, 62002334, 62121002 and U20B2047, and Students' Innovation and Entrepreneurship Foundation of USTC under Grant XY2022X01CY.

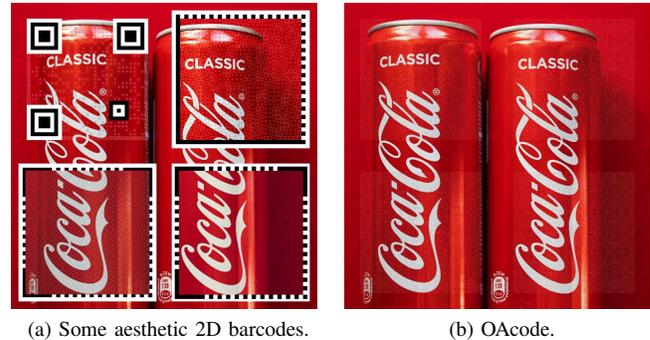


Fig. 1. Example of a commercial poster embedded with aesthetic 2D barcodes. From top left to bottom right, (a) is embedded with QR image [9], PiCode [10], RA Code [11], and RU code [12]. Their beautified data areas make the semantics of the commercial poster could be preserved, but the black and white appearance of position detection patterns and the white quiet area cause a visual discontinuity in the poster. (b) embeds four OAcodes in the corresponding positions and have a better overall visual quality.

barcodes to marketing-related materials and displaying them on various screens, potential customers can easily access the relevant information about the product by scanning these 2D barcodes with their smart mobile devices.

However, traditional barcodes, like Quick Response (QR) code [13] and Data Matrix [14], have some drawbacks in mobile marketing. Their binary appearance is initially designed to be machine-friendly for fast data transmission, yet it lacks readable hints to humans. People cannot know what information is linked to a traditional 2D barcode before scanning it, which may limit the interest of potential consumers. Additionally, their black-and-white appearance is difficult to meet the increasing demand for visual quality [15]–[17], especially in mobile marketing on the screen [6], [18]–[20]. As a result, it is hard to integrate them with colorful and well-designed mobile marketing materials.

To solve these problems, many types of aesthetic 2D barcode schemes are proposed, including the halftone 2D barcode [21], [22], the picture combination 2D barcode [23], and the picture-embedding [9]–[12], [24]–[29], where the picture-embedding 2D barcode is considered the most promising technology [27]. It generally embeds a data-related picture in the data area to generate the aesthetic 2D barcode, making them more human-readable, attractive, and better integrated with marketing materials. As shown in Fig. 1 (a), the main content of the commercial poster would not be obscured by these aesthetic 2D barcodes. Yet it could be founded that these existing schemes still retain obtrusive position detection patterns. Because in most 2D barcode detection schemes, even some state-of-art ones [30]–[32], position detection patterns

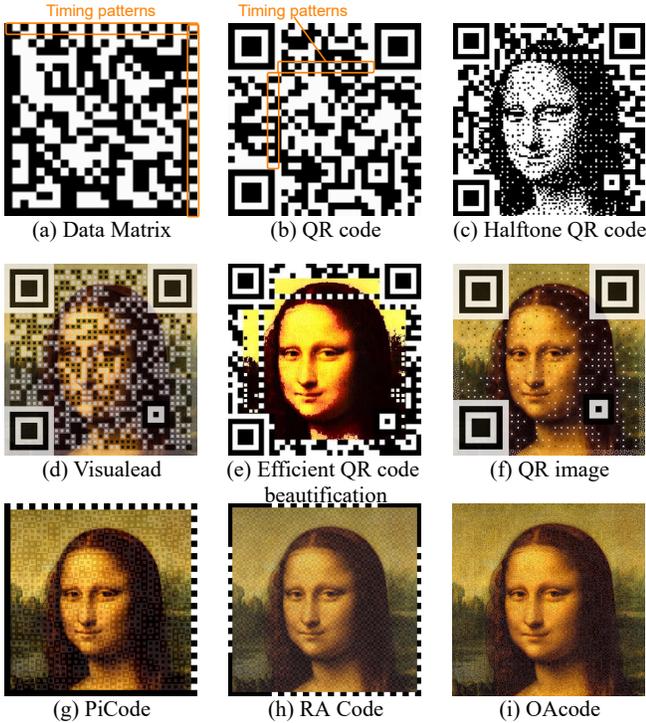


Fig. 2. Some examples of 2D barcode schemes. From the top left to the bottom right, the shown 2D barcode is getting closer to the target picture. Except for the proposed OAcode, all 2D barcodes have obtrusive position detection patterns and timing patterns.

are essential. This design results in a low structural similarity of marketing materials before and after embedding the aesthetic 2D barcode, which is one of the most commonly used quality metrics in image quality assessment [15]–[17]. Additionally, as shown in Fig. 2 (a) and (b) and Fig. 1 (a), to ensure the position detection patterns keep the pre-designed appearance, white quiet zones should be set around them [33]. Position detection patterns and the quiet zone cause an apparent boundary between the aesthetic 2D barcode and the well-designed poster and undermine the aesthetics of the poster design. Thus, we believe that these existing aesthetic 2D barcodes are partially aesthetic.

Thus, in this paper, we propose an overall aesthetic 2D barcode, named as OAcode, which eliminates the visual impact of the position detection patterns while beautifying the data area. As shown in Fig. 1 (b), the poster embedded with OAcode preserves the original design much better than existing aesthetic barcodes.

However, the position detection patterns play a significant role in the 2D barcode detection process with the following functions:

Resynchronization In the traditional 2D barcode detection process, four corner points of the 2D barcode could be detected based on position detection patterns, and then used to determine parameters to restore the perspective transformation. Moreover, position detection patterns are usually designed as rotationally invariant to determine the rotation state of the 2D barcode. Both of these are combined to resynchronize the scanned 2D barcodes to the unrotated square state.

Parameters determination To cope with the capacity requirements in different scenarios, a 2D barcode scheme usually has multiple capacity versions to choose from, which requires the extraction program to be able to determine the version of the scanned 2D barcode. In the existing traditional and aesthetic 2D barcode schemes, the version parameters are transmitted by the timing pattern, which is a part of position detection patterns as shown in Fig. 2 (a) and (b). Besides the capacity version, the extraction program needs to know the module¹ size, which could also be obtained from the timing pattern.

Lens distortion correction The lens distortion introduced by photographing process, which results in a line-to-curve mapping [34], would decrease the demodulation accuracy of 2D barcodes, especially in its large capacity version [10]. The existing 2D barcode schemes usually leverage timing patterns to resist lens distortions.

The difficulty of designing OAcode is that it, as one aesthetic 2D barcode scheme, should achieve these necessary functions mentioned above without position detection patterns. The novel aspects of OAcode, i.e., the main contributions of this paper, are summarized as follows.

- 1) OAcode cancels the position detection patterns and thus has a better visual quality. The elimination of position detection patterns is made possible by the designed symmetrical data area of OAcode, whose symmetry is used as synchronizing signal in the following detection process. In the OAcode generation process, we also propose a fast blending method with adaptive intensity, which can play a similar role to error diffusion dithering and avoid the clipped intensity due to the content of the background image.
- 2) We carefully design an OAcode detection strategy based on its data area symmetry and Hough transform, which could implement all functions that should have been realized by visible position detection patterns, including 2D barcode synchronization, rotation state determination, module size determination, and information capacity determination.
- 3) We propose an enhanced demodulation method for OAcode to obtain a higher demodulation accuracy under the lens distortion, which commonly exists in the camera-shooting process [34]–[36]. Specifically, the result of the first demodulation, i.e., the accumulated data unit would be divided into subunits to execute the seconding position detection process as reference signals and re-accumulate the detection results to obtain a new data unit with less lens distortion.

II. RELATED WORK

The proposed OAcode is a kind of aesthetic 2D barcode scheme, which embeds information into one image. Thus, in this section, we first review some related aesthetic 2D barcode schemes and then discuss the difference between aesthetic 2D barcodes and some information embedding techniques.

A. Aesthetic 2D Barcode

Compared with traditional 2D barcodes (Fig. 2 (a) and (b)) with an obtrusive appearance, aesthetic 2D barcodes (Fig. 2

¹According to the barcode terminology, the smallest square unit with the black-and-white appearance used to represent bit ‘0’ or ‘1’ is called a module.

(c)~(i)) are human-readable, more attractive, and can be better integrated with marketing materials. In addition, almost all aesthetic 2D barcode schemes have the same design, i.e., the combination of position detection patterns and beautified data area, in which their position detection patterns have a similar appearance and the same functionality as those of traditional 2D barcodes. For example, we could find that the structure of position detection patterns in Fig. 2 (c)~(h) is similar to ones shown in Fig. 2 (a) and (b).

The existing aesthetic 2D barcode schemes could be roughly divided into two categories according to their design purposes. The first category is generated by modifying an existing traditional 2D barcode scheme. These schemes are compatible with the standard extractor and have better visual quality. Because of the widespread usage of QR code, most schemes in this category are beautified QR code schemes [9], [22]–[27]. Fig. 2 (c)~(f) show some examples. These schemes usually modify the generation process of QR code, mainly the modulation process, to improve the visual quality. For compatibility, they retain the same extraction process as QR code, resulting in the same position detection patterns.

And another category includes the aesthetic 2D barcode schemes [10]–[12], [28], [29] generated independently of existing black-white 2D barcodes. Schemes in this category need not be compatible with existing 2D barcode extractors. Therefore they usually have more design freedom and better performance compared with schemes in the above category. As shown in Fig. 2 (g) and (h), one obvious visual improvement of aesthetic 2D barcodes in this category is that they could design smaller position detection patterns, which usually means a closer appearance to the background image and better visual quality.

The proposed OAcodes belongs to the second category, yet compared to the previous schemes, OAcodes has a completely different design, i.e., the elimination of position detection patterns. As a result, OAcodes is visually closer to one image (Fig. 2 (i)) and can be better integrated with one poster (Fig. 1 (b)).

B. Information Embedding Techniques

We discuss the following information embedding techniques, including watermark, steganography, and deep data hiding. Watermark and steganography are the two main branches of information hiding [37], in which digital image watermarking [36], [38]–[40] is frequently used to prove image ownership, designed with a focus on the visual imperceptibility and the robustness to various digital image distortions. Steganography [41]–[43] is used for covert communication, so the corresponding methods focus on the security and the capacity of the embedded secret message, in which the security usually means the hidden message is imperceptible to steganalysts. Deep data hiding [44]–[47] is an emerging category of techniques that use neural networks to hide information in images and extract it through camera shooting. Deep data hiding focuses on the robustness of camera shooting distortions, which are more serious than digital image distortions.

Although OAcodes has a similar appearance to other information embedding techniques mentioned above, it follows the

TABLE I:

The different design criteria of information embedding techniques and aesthetic 2D barcode.

Picture-embedding methods	Robustness	Imperceptibility	Capacity
Steganography	low	high	high
Watermarking	medium	medium	low
Deep data hiding	high	medium	medium
Aesthetic 2D barcode	high	low	high

same design criteria as other aesthetic 2D barcode schemes, which makes OAcodes different from these information embedding techniques. Specifically, one 2D barcode scheme always uses modules to represent bits ‘0’ or ‘1’. For example, the modules of QR code are black-and-white square units and OAcodes uses the spread spectrum module. Then, in the blending process, the background image only affects the blending intensity of these modules. Based on modules, 2D barcode schemes could design a parameter determination strategy, which allows its detector to extract one 2D barcode with unknown parameters, e.g., unknown capacity. Considering the different capacity and robustness requirements of 2D barcodes in various application scenarios, this strategy allows these 2D barcodes with different settings could be decoded by the same detector. While for most information embedding techniques mentioned above, the image content and the embedded information together determine the modification intensity and patterns on the host image to better balance the robustness and imperceptibility. Additionally, their end-to-end design has excellent performance with preset parameters but is not suitable for input with variable parameters.

As summarized in Table I, these four research areas are designed for different applications and have different concerns and design criteria.

III. SYMMETRY CALCULATION

The proposed OAcodes has a novel synchronization process, in which its symmetry is used as the calibration signal. To fast calculate symmetry, we slightly modify the method proposed by Ma et al. [48]. In this section, we first define the symmetry of a 2D matrix and then illustrate its calculation method.

1) *Symmetry Definition:* If 2D matrix A is symmetrical about point (x_s, y_s) , A satisfies:

$$A(x_s - x, y_s - y) = A(x_s + x, y_s + y). \quad (1)$$

The point (x_s, y_s) is one symmetrical center of A . The symmetry S of A about a point (i, j) can be defined by:

$$S(i, j) = \sum_x \sum_y A(i - x, j - y)A(i + x, j + y). \quad (2)$$

$S(i, j)$ is actually the correlation between A and the flipped A around point (i, j) . Meanwhile, given a point (i, j) , the value of $S(i, j)$ represents the possibility that point (i, j) is one of the symmetrical centers of A . In the ideal case, the positions of symmetrical peaks in S are the positions of symmetrical centers in A .

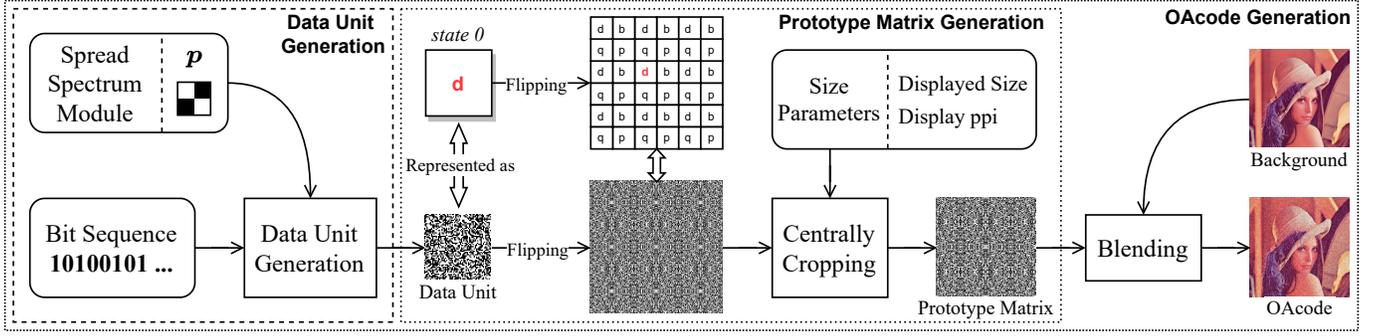


Fig. 3. The generation process of OAcoding, including data unit generation, prototype matrix generation, and blending process.

2) *Auto-convolution Function*: Obviously, calculating symmetry in space domain like Eq. (2) is inefficient, especially when the A has a large size. To decrease the computational cost of Eq. (2), it is first rewritten as:

$$\begin{aligned} S(i, j) &= \sum_x \sum_y A(i-x, j-y)A(i+x, j+y) \\ &= \sum_x \sum_y A(x, y)A(2i-x, 2j-y), \end{aligned} \quad (3)$$

where x and y run over all values that lead to legal subscripts of A . Defining a temporary matrix T , $T(2i, 2j) = S(i, j)$ and substitute it to Eq. (3):

$$\begin{aligned} T(2i, 2j) &= \sum_x \sum_y A(x, y)A(2i-x, 2j-y) \\ T(u, v) &= \sum_x \sum_y A_p(x, y)A_p(u-x, v-y), \end{aligned} \quad (4)$$

where A_p is A zero-padding to doubly the original size. It could be observed that T is the auto-convolution of A_p . Thus, the convolution theorem could be used to get the frequency form of Eq. (4) as follows:

$$T = IFFT[FFT(A_p)FFT(A_p)], \quad (5)$$

where FFT represents the fast Fourier transform and $IFFT$ is the corresponding inverse transform. Considering T is the doubly upsampling of S , the symmetry S of 2D matrix A can be calculated by:

$$S = \mathcal{D}(IFFT[FFT(A_p)FFT(A_p)]), \quad (6)$$

where $\mathcal{D}(\cdot)$ is a downsampling function scaling its input matrix to half size. Using the frequency form of auto-convolution function, the computational cost of calculating symmetry S is greatly reduced from Eq. (2) to Eq. (6).

It should be noted that compared with [48], OAcoding proposes a comprehensive restoration strategy for the perspective distortion in case of large information capacity, which is optimized for the 2D barcode scanning process, including the estimation of perspective transformation in the absence of symmetric peaks (Sec. V-B) and the demodulation enhancement scheme for lens distortion. (Sec. V-D2).

IV. OACODING GENERATION

Fig. 3 is the flowchart of the generation process of OAcoding. The input message is converted to the bit sequence to generate the data unit. Then, the data unit is flipped to generate one symmetrical matrix, which would be further cropped according to a set of size parameters to generate the prototype matrix. Finally, we blend the prototype matrix and the background image to obtain the OAcoding.

A. Data Unit Generation

The input message is converted to a bit sequence and reshaped to a 2D bit matrix m of size $L_m \times L_m$. Then, m is spread spectrum encoded using p to generate a data matrix D , where p is the spread spectrum module, a bipolar matrix with ± 1 elements alternately arranged like a chessboard. In the spread spectrum process, bit '1' in the 2D bit matrix m is presented as $+1 \times p$ in the data matrix D , and bit '0' in m is represented as $-1 \times p$ in D . According to the spread spectrum theory, transmitting over a larger bandwidth could increase the robustness against external narrowband interference [49], [50]. Thus, we could use a larger spread spectrum module to obtain better robustness against imaging distortions, e.g., blur, noise, and compression, in the screen-camera process. Finally, a random bipolar matrix is used as the mask matrix K to mask D to generate the data unit U :

$$U(i, j) = D(i, j)K(i, j), \quad (7)$$

where i and j denote the index of these matrices. The masking operation makes the distribution of black and white pixels in data unit U more random, which is beneficial to the OAcoding extraction. Additionally, the mask matrix K and the spread spectrum module p would be leveraged to estimate the parameters needed in the extraction process, such as the information capacity of OAcoding.

B. Prototype Matrix Generation

The unrotated and unflipped state of the data unit U is defined as the initial state *state 0*. Then, data U with *state 0* is flipped to generate one symmetrical matrix with many symmetrical centers, which are located on the corners of data units. To better illustrate the flipping process, U with *state 0* is represented as a block with a red and bold symbol 'd'; see

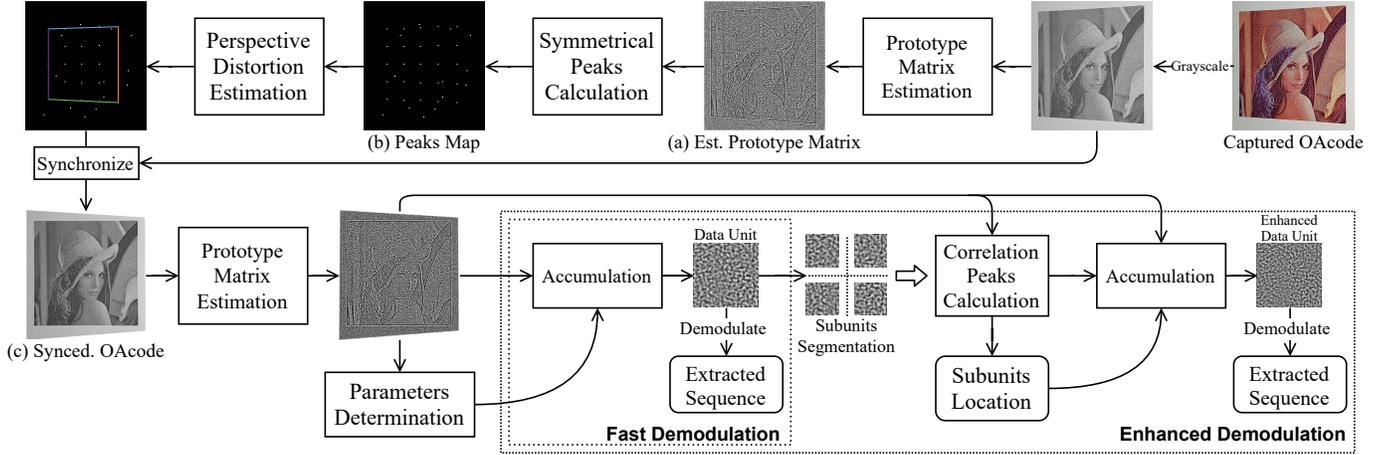


Fig. 4. The extraction process of OAcoding. The input captured OAcoding is used to estimate the prototype matrix in (a), whose symmetrical peaks are calculated to generate the peaks map (b). Then, based on the peaks map, parameters of the perspective distortion could be estimated to synchronize OAcoding. Finally, implementing two-stage demodulation on the synchronized OAcoding (c) to extract the message sequence.

Fig. 3. After the flipping process, that data unit is positioned on the upper left of the global symmetrical center. This positional relationship would be used in the parameter determination strategy of the OAcoding extraction process.

Then, the symmetrical matrix is centrally cropped to generate the prototype matrix with the target size, which is determined by two size parameters, including the pixel per inch (PPI) p of the display and the target displayed size L_d . Specifically, the cropping size L_c can be calculated by: $L_c = (L_d \times p) / C_f$, where $C_f = 2.54$, is the conversion factor for inches and centimeters. For example, if we need to display an OAcoding of size 5×5 cm in iPad mini 2 with PPI $p = 324$, substituting the above parameters, we can obtain the cropping size $L_c = 638$, which means the OAcoding should be 638×638 pixels size to achieve the target display effect.

A noteworthy issue is that the determined size of the prototype matrix would restrict the available information capacity, which is a reasonable result because we cannot display one accessible 2D barcode with a large information capacity in one limited region. In the case of OAcoding, a larger information capacity means a data unit with a larger size, resulting in fewer flipped data units in the prototype matrix with the fixed size L_c . Yet for better OAcoding detection, there should be enough symmetrical centers, i.e., enough data units in OAcoding. The size parameters and the information capacity of OAcoding restrict each other. In this paper, for better extraction, the parameters we selected make the prototype matrix P contain 5×5 or more symmetrical centers, i.e., more than 5×5 data units in P .

C. Blending

Now we have a bipolar prototype matrix P , and it is blended with the input background image I to generate the beautified P , i.e., the OAcoding O . Assuming the background image I is an 8-bit grayscale image, whose values range from 0 to 255 and have been resized to the same size as P . The proposed

OAcoding O could be simply generated by:

$$O(i, j) = I(i, j) + \lambda P(i, j), \quad (8)$$

where λ is the blending intensity. But considering the clipped intensity, OAcoding generated by Eq. (8) would be weakened while the background image has brighter or darker contents. To solve the above problem and balance the visual quality and extraction performance, a modulation method with adaptive intensity is proposed:

$$\begin{aligned} O(i, j) &= I(i, j) + \lambda P(i, j) - 2\lambda(I(i, j)/255 - 0.5) \\ &= I(i, j) + \lambda \left(P(i, j) + 1 - \frac{2}{255} I(i, j) \right). \end{aligned} \quad (9)$$

OAcoding generated by Eq. (9) is brighter in the darker image content and darker in the brighter content, which is similar to the result of error diffusion dithering. Moreover, Eq. (9) could be calculated by matrix operations with a low computational cost. In the case of the colorful background image, the input image is first converted from RGB channels to $YCbCr$ channels, and its Y component would be taken as I to execute the above operations. Finally, the grayscale OAcoding is combined with Cb, Cr components to convert back to a colorful OAcoding.

V. OACODING EXTRACTION

Fig. 4 illustrates the extraction process of the proposed OAcoding, in which we assume the OAcoding captured by the camera is the only input. To synchronize the captured OAcoding, we first estimate the prototype matrix and calculate its symmetrical peaks, which represent the position of the data unit corners and could be used to estimate the geometric distortions, mainly the perspective distortion, occurring on OAcoding. After the restoration of perspective distortion, the same as other 2D barcode schemes, we estimate the necessary parameters to demodulate OAcoding, including rotation state, information capacity version, and module size. Based on these parameters, OAcoding could be demodulated to a bit sequence.

Additionally, to resist lens distortion, we propose enhanced demodulation to further improve extraction accuracy.

A. Symmetrical Peaks Calculation

In this process, taking the captured OAcodes as the input, we calculate its symmetrical peaks map \mathbf{M} for the following OAcodes synchronization.

The captured OAcodes is first converted to grayscale OAcodes \mathbf{O}' . Then, to decrease the impact of the background image, we separate the prototype matrix \mathbf{P} from OAcodes \mathbf{O}' according to the near-noise appearance of the prototype matrix. Specifically, the prototype matrix \mathbf{P} could be estimated by filtering according to [48], [51]:

$$\widehat{\mathbf{P}} = (\mathbf{O}' - \boldsymbol{\mu}) \frac{\sigma_{\mathbf{P}}^2}{\sigma_{\mathbf{O}'}^2}, \quad (10)$$

where $\widehat{\mathbf{P}}$ is the estimated prototype matrix (Fig. 4 (a)), $\boldsymbol{\mu}$ is the local mean of \mathbf{O}' , and $\sigma_{\mathbf{P}}^2$ and $\sigma_{\mathbf{O}'}^2$ are the local variance of \mathbf{P} and \mathbf{O}' . It should be noted that $\sigma_{\mathbf{P}}^2$ is a constant matrix whose values mainly depend on the intensity λ in Eq. (8). Thus, Eq. (10) could be simplified as:

$$\widehat{\mathbf{P}} = \frac{\mathbf{O}' - \boldsymbol{\mu}}{\sigma_{\mathbf{O}'}^2}. \quad (11)$$

Then, the symmetry matrix \mathbf{S} of $\widehat{\mathbf{P}}$ can be calculated by the frequency form of the autoconvolution function mentioned in Sec. III-2 as follows:

$$\mathbf{S} = \mathcal{D}(\text{IFFT}[\text{FFT}(\widehat{\mathbf{P}}_p)\text{FFT}(\widehat{\mathbf{P}}_p)]), \quad (12)$$

where $\widehat{\mathbf{P}}_p$ is $\widehat{\mathbf{P}}$ zero-padding to doubly the origin size and $\mathcal{D}(\cdot)$ is a downsampling function which scales its input matrix to the half size.

Finally, an adaptive symmetrical peaks filter is adopted to generate the symmetrical peaks map as follows:

$$\mathbf{M} = \begin{cases} 1 & , \text{ if } S > \mu_S + \beta\sigma_S^2 \\ 0 & , \text{ otherwise} \end{cases} \quad (13)$$

where μ_S and σ_S^2 respectively denote the local mean and variance of \mathbf{S} . The coefficient β is an adjustable parameter controlling the number of filtered symmetrical peaks. As mentioned in Sec. IV-B, there should be more than 25 symmetrical centers in the prototype matrix \mathbf{P} . Thus, as shown in Fig. 4 (b), considering potential noise peaks, we adjust β until the peaks map \mathbf{M} has slightly more than the expected 25 peaks.

B. Perspective Distortion Restoration

In this process, we estimate perspective distortion parameters from symmetrical peaks map \mathbf{M} and apply the corresponding restoration to the captured OAcodes.

For the peaks map \mathbf{M} , the first technical question is: *How to distinguish symmetrical peaks and noise peaks?* Assuming that the peaks generated by noise are randomly distributed, we can take advantage of the regular distribution of symmetrical peaks, which comes from the grid-type distribution of data units, to distinguish symmetrical peaks. Specifically, these

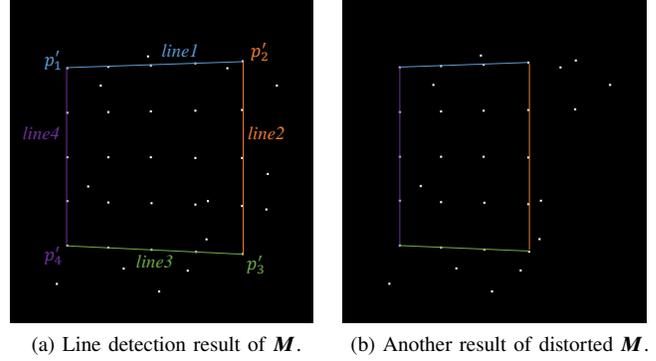


Fig. 5. Two different straight line detection results of similar peaks maps \mathbf{M} s. The peak map in (b) misses some peaks on the right bottom, resulting in the detected right straight line being closer to the center.

symmetrical peaks represent the corners of data units, so they are also on the edges of data units and are distributed on different straight lines in groups. Thus, we use Hough transform to detect straight lines, which pass through these symmetrical peaks. As shown in Fig. 5 (a), in all detected lines, we select four outermost lines *line1*, *line2*, *line3*, *line4* and calculate their intersections p'_1 , p'_2 , p'_3 , p'_4 , which constitute the known points set \mathbf{P}' . To restore perspective distortion, we also need the target points set \mathbf{P} , whose points p_1 , p_2 , p_3 , p_4 have the undistorted original coordinates of points in \mathbf{P}' .

In the ideal case, points in known points set \mathbf{P}' are the outermost four data unit corners of OAcodes, whose relative positions are easily determined. However, considering the serious distortion in the camera-shooting process, sometimes a few peaks are missed in \mathbf{M} . Fig. 5 (b) shows another possible line detection result when missing some symmetrical peaks. Thus, the second technical question is: *How to determine the target point set \mathbf{P} under different line detection results?* For example, Fig. 5 (a) and (b) have two different target point set \mathbf{P} . Our solution takes advantage of the fact that data units are equally spaced, which means the distance between every two symmetrical peaks on the same line is similar even after perspective distortion. We record all distances between every two adjacent points on *line1*~*line4* and use K-means clustering to obtain the biggest distance cluster, whose average value is taken as the estimated side length of the data unit, denoted as L_u . Denoting the length of *line1*~*line4* as L_1 ~ L_4 , the target point set \mathbf{P} can be calculated by:

$$\begin{aligned} p_1(x_1, y_1) &= (x_1, y_1) \\ p_2(x_2, y_2) &= \left(x_1 + \mathcal{R}\left(\frac{L_1 + L_3}{2L_u}\right)L_u, y_1 \right) \\ p_3(x_3, y_3) &= \left(x_1 + \mathcal{R}\left(\frac{L_1 + L_3}{2L_u}\right)L_u, y_1 + \mathcal{R}\left(\frac{L_2 + L_4}{2L_u}\right)L_u \right) \\ p_4(x_4, y_4) &= \left(x_1, y_1 + \mathcal{R}\left(\frac{L_2 + L_4}{2L_u}\right)L_u \right), \end{aligned} \quad (14)$$

where $\mathcal{R}(\cdot)$ is the rounding function. There is no specific constraint to the coordinates of p_1 , and we make $p_1 = p'_1$

empirically. Finally, substituting all coordinates (x', y') and (x, y) of the point in P' and P into the following equations:

$$\begin{aligned} x &= \frac{a_1x' + b_1y' + c_1}{a_0x' + b_0y' + 1} \\ y &= \frac{a_2x' + b_2y' + c_2}{a_0x' + b_0y' + 1}, \end{aligned} \quad (15)$$

we can solve these equations to obtain all needed parameters $a_0, b_0, a_1, b_1, c_1, a_2, b_2, c_2$ for the perspective distortion restoration. As shown in Fig. 4 (c), the captured OAcodes would be synchronized to the square state.

C. Parameters Determination

To demodulate the restored OAcodes, three parameters should be determined, including the rotation state, the information capacity, and the spread spectrum module. In most traditional and aesthetic 2D barcode schemes, these parameters are detected from visible position detection patterns. For the proposed OAcodes without position detection patterns, we design a parameter determination strategy based on the statistical characteristics of OAcodes.

First, we select a specific data unit in OAcodes for analysis. The synchronized OAcodes in the previous subsection is used to estimate a synchronized prototype matrix following the same process mentioned in Sec. V-A, which is denoted as \hat{P} for notational simplicity. Without the perspective distortion, the symmetry of the synchronized prototype matrix \hat{P} would be significantly enhanced. Therefore it is reasonable to assume that all the symmetrical peaks are detected by Eq. (12) and Eq. (13), i.e., all data units U_s in \hat{P} could be located. Then, the data unit on the upper left of the global symmetrical center G is selected as the data unit to be analyzed, denoted as U_a .

According to the setting in Sec. IV-B, there are two possible states of U_a , as shown in Fig. 6. When OAcodes is not rotated or rotated 180° , U_a would be the original state, *state 0*. If OAcodes is rotated 90° or 270° , U_a would be the *state 1*. The information capacity and size of the spread spectrum matrix p depend on the setting in the OAcodes generation process. In this paper, OAcodes has two capacity settings, denoted by *capacity 0* and *capacity 1*, and three types of spread spectrum module size, denoted by *module 0*, *module 1*, and *module 2*.

For all possible *state i*, *capacity j*, and *module k*, the confidence C_{ijk} that these parameters are correct could be calculated by following steps. According to *state i*, U_a is first rotated and flipped back to *state 0*. Then, U_a is scaled to an expected size determined by *capacity j* and *module k*. For example, in the case of *capacity 0* and *module 0*, the data unit is generated with $L_m \times L_m$ spread spectrum modules ps , where $L_m = 32$ and $p = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}$. Thus, the expected size of U_a is 64×64 . Then, we use the same mask matrix K as Sec. IV-A to demask U_a to obtain the corresponding data matrix D_a :

$$D_a(x, y) = U_a(x, y)K(x, y), \quad (16)$$

where x and y denote the index of these matrices. If all parameters are correct, data matrix D_a would be a matrix composed of spread spectrum matrix p and $-p$. But if any parameter is wrong, D_a would be the result of a wrong

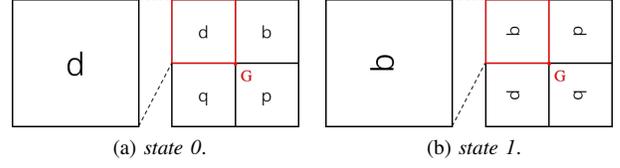


Fig. 6. Two possible states of the analyzed data unit U_a , i.e., the data unit on the upper left of the global symmetrical center G of the OAcodes.

demasking, close to a random matrix. We divide D_a into $L_m \times L_m$ non-overlapping blocks with the same size as p . Denoting the block as $D_a^{x,y}$, the confidence C_{ijk} is calculated by:

$$C_{ijk} = \sum_x \sum_y \left| \langle \bar{D}_a^{x,y}, p \rangle_{\mathbf{F}} \right|, \quad (17)$$

where $\bar{D}_a^{x,y}$ is the normalized $D_a^{x,y}$ and $\langle \cdot, \cdot \rangle_{\mathbf{F}}$ is the Frobenius inner product: a component-wise inner product of two matrices. Finally, the *state i*, *capacity j* and *module k* of U_a can be determined by:

$$i, j, k = \arg \max_{i,j,k} C_{ijk}. \quad (18)$$

The rest data units have the same *capacity* and *module* with U_a , and their *states* can be easily determined according to the relative position of themselves and U_a .

D. Demodulation

We propose a two stages demodulation method to meet different requirements of accuracy and computational cost in the demodulation process of OAcodes. The first stage conducts a fast demodulation method with less computational cost. In most cases, its demodulation result is accurate enough. When users pursue higher demodulation accuracy, enhanced demodulation in the second stage would be adopted, which considerably decreases demodulation error caused by lens distortion.

1) *Fast Demodulation*: According to the location and parameters of data units determined in Sec. V-C, we restore all data units back to *state 0* and accumulate them to generate one accumulated data unit \hat{U} . The accumulation operation can efficiently reduce the impact of random noise. Similar to Sec. V-C, we resize the accumulated data unit \hat{U} to the expected size according to its *capacity* and *module*. Then, the accumulated data unit \hat{U} is demasked with the mask matrix K to obtain the data matrix \hat{D} , which is divided into $L_m \times L_m$ non-overlapping and normalized blocks $\hat{D}^{x,y}$. To despread data matrix \hat{D} , the correlation $\rho_{x,y}$ between $\hat{D}^{x,y}$ and p is calculated by:

$$\rho_{x,y} = \langle \hat{D}^{x,y}, p \rangle_{\mathbf{F}}, \quad (19)$$

where $\langle \cdot, \cdot \rangle_{\mathbf{F}}$ is the Frobenius inner product, x and y represent the block index in \hat{D} , and the spread spectrum module p is determined by the *module* parameter. Then, the

message matrix \mathbf{m}' can be extracted by:

$$m'_{x,y} = \begin{cases} 0 & , \rho_{x,y} < 0 \\ 1 & , \rho_{x,y} \geq 0 \end{cases} \quad (20)$$

Finally, we reshape the message matrix to a bit sequence, i.e., the extracted message.

2) *Enhanced Demodulation*: Considering the imperceptible lens distortion in the captured OAcodes, the accuracy of the fast demodulation could be further improved by the enhanced demodulation, which reduces the impact of lens distortions. We first discuss how lens distortions impact the demodulation accuracy and then illustrate the implementation of the enhanced demodulation method.

In fast demodulation, all data units in OAcodes are accumulated to obtain one accumulated data unit \hat{U} to enhance the meaningful signal and weaken random noise. However, the lens distortion in different data units is different, which causes misalignments in the data unit accumulation and weakens the accumulation result. Ignoring other factors, the smaller the size of one block in the camera-shooting OAcodes, the weaker the lens distortion, and the less the bit error rate caused by misalignments. Inspired by this rule, we design the following enhanced demodulation method.

In this enhanced demodulation, the steps up to the generation of the accumulated data unit \hat{U} are the same as the fast demodulation method. Then, \hat{U} is used for the second position detection of data units rather than direct demodulation. The first position detection of data units is in Sec. V-A, where the corners of data units are detected as the symmetrical peaks by the auto-convolution function. The second position detection of data units is based on the correlation peaks calculated by the cross-correlation between the accumulated data unit \hat{U} and the synchronized prototype matrix \hat{P} . To decrease the impact of lens distortions, we take the smaller subunit to execute the second position detection. Specifically, as shown in Fig. 7 (a), we divide the accumulated data unit \hat{U} into $n \times n$ non-overlapping subunits $\hat{U}^{x,y}$. Denoting L_x and W_y as the length and the width of the subunit $\hat{U}^{x,y}$, and L_U as the size of the accumulated data unit \hat{U} , L_x and W_y can be determined by:

$$L_x = \begin{cases} \lfloor \frac{L_U}{n} \rfloor & , x \in [1, n-1] \\ L_U - (n-1) \lfloor \frac{L_U}{n} \rfloor & , x = n \end{cases} \quad (21)$$

$$W_y = \begin{cases} \lfloor \frac{L_U}{n} \rfloor & , y \in [1, n-1] \\ L_U - (n-1) \lfloor \frac{L_U}{n} \rfloor & , y = n \end{cases}$$

Then the cross-correlation C_{xy} between subunit $\hat{U}^{x,y}$ and the synchronized prototype matrix \hat{P} can be calculated by:

$$C_{xy} = \text{IFFT}[\text{FFT}(\hat{U}^{x,y})\text{FFT}(\hat{P})^*] \quad (22)$$

where $*$ represents the conjugate of the matrix. As shown in Fig. 7 (b), according to correlation peaks in C_{xy} , we could locate the corresponding subunits and accumulate them to generate an enhanced subunit. Repeating the above operations for all possible x and y , an enhanced data unit with less lens distortion can be obtained. As shown in Fig. 7 (c) and (a), the enhanced data unit is much clearer and has a much lower bit error rate (1.43%) than the original one (11.18%). In the

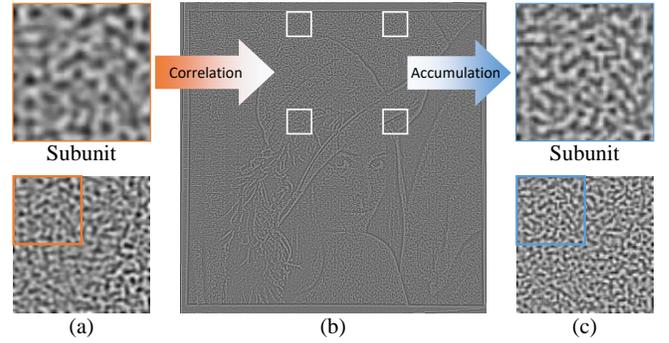


Fig. 7. An example of the second position detection of data units. (a) is the accumulated data unit, which is divided into 4 subunits. (b) is the synchronized prototype matrix, in which white borders are the detection results of one subunit. (c) is the enhanced data unit, generated by the accumulation of these detected subunits.

practice, the enhanced subunit is accumulated by only four subunits with the largest four correlation peaks, because the weak correlation peak means an unclear subunit, which brings more interference than improvement. Finally, the enhanced data unit is demodulated by the same process as the one in Sec. V-D1.

VI. EXPERIMENT EVALUATION

A. Experimental Methodology

In this subsection, we illustrate the default experimental setting. For a specific experiment, except for some rewritten parameters, most experimental parameters are the same as the default setting.

In the OAcodes generation process, the prototype matrix is set to 32×32 (1024) bits information capacity and modulated by 3×3 size spread spectrum module \mathbf{p} . Then, the prototype matrix is cropped to 640×640 pixels and blended with the background image to generate the tested OAcodes, in which the blending intensity λ is set to 20.

In the extraction process, two screens are used to display OAcodes. One is iPad mini 2, whose resolution is 2048×1536 pixels and PPI (pixel per inch) is 324. And the other is Huawei P30 Pro, having a 2340×1080 pixels resolution and 398 PPI. For the OAcodes with the default size setting of 640×640 pixels, it would be displayed with a size of $5 \times 5\text{cm}$ on iPad mini 2 and a size of $4 \times 4\text{cm}$ on Huawei P30 Pro. The illuminance of displays is set to $80 \pm 10\text{lux}$. Then, OAcodes are captured by two mobile phones, iPhone 11 and Xiaomi Mi 6 with a limited capturing resolution of 720p (1280×720 pixels) at the screen-camera distance of 10cm . The camera-shooting process is done indoors with an illuminance of around 300lux . In the enhanced demodulation, if not specified, the accumulated data unit is divided into 2×2 subunits for the second position detection.

It should be noted that the following performance of OAcodes is evaluated with clearly defined experimental parameters so that for new devices or new parameters, some performance of OAcodes on them could be reasonably inferred, i.e., the available distance.

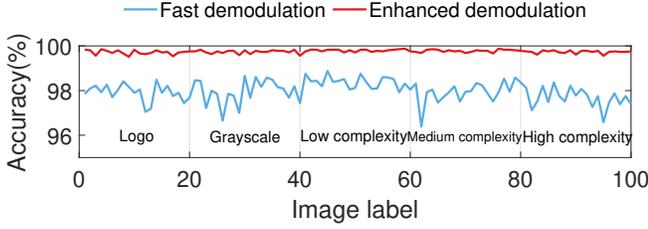


Fig. 8. Overall performance of OAcodes.

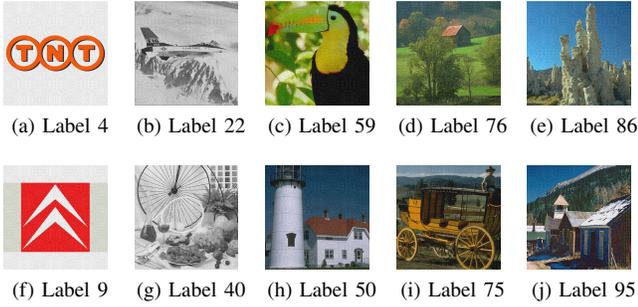


Fig. 9. Top row: OAcodes with the highest accuracy in five image categories. Bottom row: OAcodes with the lowest accuracy.

B. OAcodes Performance

1) *Overall Performance*: We select 100 background images to evaluate the overall performance of the proposed OAcodes. To cover most scenarios, the selected background images include 5 categories, logo, grayscale, low complexity color images, medium complexity color images, and high complexity color images. Each category has 20 images. The logo images are selected from BelgaLogos Dataset [52], and the other images come from the CVG-UGR Image database [53]. In this paper, we use the multiscale structural similarity index (MS-SSIM) and the peak signal-to-noise ratio (PSNR) to measure the objective visual quality of OAcodes. In this subsection, these OAcodes have an average MS-SSIM of 0.821 and an average PSNR of 22.45. Additionally, we invite 20 participants to evaluate the visual quality of OAcodes with the criteria that the score of the black-and-white QR code and the original background image is 1 and 10 respectively. The proposed OAcodes obtain a mean opinion score of 8.15, which is beneficial from its noise-like data area without visible position detection patterns.

The extraction results are shown in Fig. 8, in which the demodulation accuracy of one OAcodes is the average accuracy extracted under four possible screen-camera combinations. The blue line represents the accuracy calculated by fast demodulation, whose minimum is over 96%. The red line is the demodulation result of the enhanced demodulation with the minimum exceeding 99.5%. These extraction results illustrate that the proposed OAcodes have a very high extraction accuracy on various background images of different categories.

The images with the lowest and the highest accuracy in each category are selected as representative background images. As shown in Fig. 9, OAcodes generated by these 10 representative background images are evaluated in the subsequent experi-

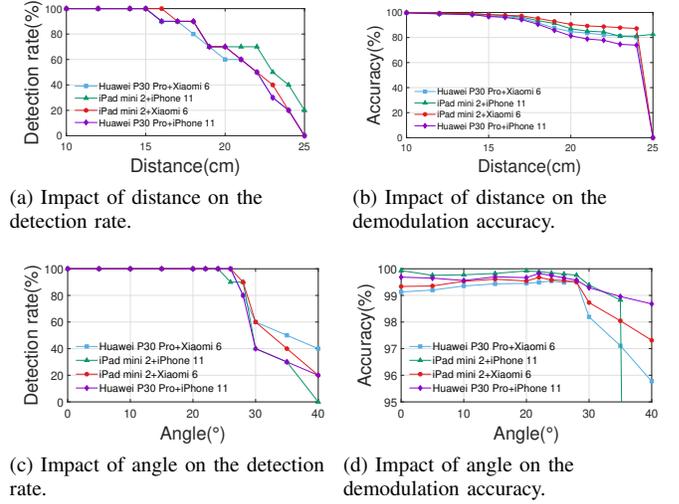


Fig. 10. The detection and demodulation range of OAcodes.

ments.

2) *Detection and Demodulation Range*: To evaluate the performance of OAcodes under different camera-shooting conditions, 10 OAcodes in Fig. 9 are captured at various screen-camera distances and angles. The extraction results are shown in Fig. 10, in which (a) and (c) illustrate the detection rate of the captured OAcodes as well as (b) and (d) illustrate the demodulation accuracy. We take the detection rate to represent the probability that the captured OAcodes could be resynchronized. Then, the synchronized OAcodes are demodulated to obtain the extracted message, whose bit recovery rate is used as the demodulation accuracy.

According to Fig. 10 (a) and (b), at 15cm screen-camera distance, OAcodes have a 100% detection rate and up to 96.76% demodulation accuracy. Although the detection rate of OAcodes drops to 60% at a screen-camera distance of 20cm, its demodulation accuracy still exceeds 80%. Considering using several frames to extract messages is common in 2D barcode applications, the detection rate of OAcodes could be theoretically improved to 93.6% using 3 frames at a distance of 20cm. Because of the high demodulation accuracy of OAcodes, OAcodes could be extracted at a distance of 24cm by sufficiently scanning. In the angle experiment, the screen-camera distance is fixed as 10cm. As shown in Fig. 10 (c) and (d), at a 25° screen-camera angle, OAcodes could be detected and demodulated almost without error. Similarly, at a scanning angle of 35°, OAcodes can be theoretically extracted with over 97% demodulation accuracy.

3) *Impact of Spread Spectrum Module Size*: The size of the spread spectrum module \mathbf{p} , $\mathbf{p_size}$, has a considerable influence on the performance of OAcodes. According to the interconstraint relationship between parameters mentioned in Sec. IV-B, when OAcodes is set to 640×640 pixels and 1024 bits information capacity, the size of \mathbf{p} can only be 2×2 , 3×3 , or 4×4 pixels, corresponding to three types of OAcodes. Then, based on 10 representative background images, we generate 3×10 OAcodes and evaluate their performance differences caused by $\mathbf{p_size}$.

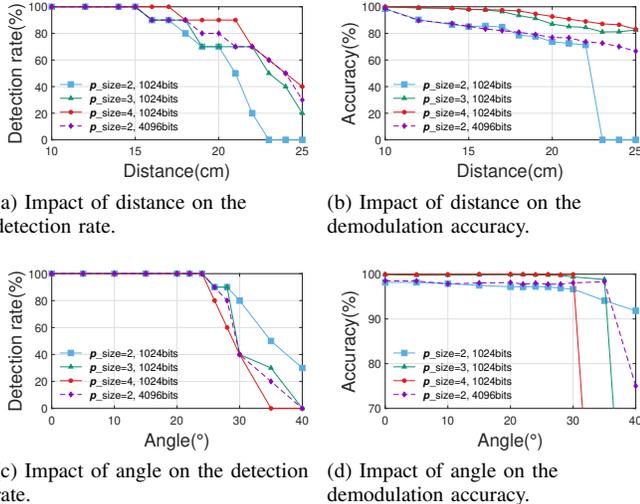


Fig. 11. The detection and demodulation range of OAcodes generated by different spread spectrum modules p .

As shown in Fig. 11, blue, green, and red lines illustrate the detection and demodulation range of OAcodes with spread spectrum module size $p_size = 2, 3$, and 4 pixels. When $p_size = 4$, OAcodes have the best demodulation accuracy in both distance and angle experiments (83.13% at 25cm and 99.98% at 30°). Moreover, it has a better detection rate in the distance experiment (40% at 25cm). These results are consistent with the conclusion mentioned in Sec. IV-A, i.e., the larger size of spread spectrum module p contributes to better robustness against imaging distortions.

On the other hand, the bigger spread spectrum module size p_size means fewer data units in OAcodes, i.e., fewer symmetrical peaks in peaks map, which increase the difficulty of OAcodes synchronization. This explains why OAcodes with $p_size = 4$ has the lowest detection rate in the angle experiment, as shown in Fig. 11 (c). On the contrary, the smaller p means better synchronization performance and lower demodulation accuracy. As shown in Fig. 11 (c), when $p_size = 2$, OAcodes have a relatively better detection rate 30% at angle 30°. Yet it has a relatively lower demodulation accuracy in Fig. 11 (b) and (d). For the OAcodes generated by 3×3 size spread spectrum module, it has an intermediate detection and demodulation performance between the above two.

4) *Maximum Information Capacity*: The upper bound of the information capacity of OAcodes can be determined according to the discussion of parameters in Sec. IV-B. Specifically, there should be more than 5×5 data units in one OAcodes, and the spread spectrum module p should be larger than or equal to 2×2 pixels. Thus, the maximum bit pixel rate of OAcodes is $1/100$. For OAcodes of 640×640 pixels in experiments, its maximum information capacity is 4096 bits, whose performance is shown in Fig. 11 as the purple dotted line. At an 18cm screen-camera distance, that OAcodes have a 90% detection rate and over 80% demodulation accuracy. Additionally, at a 25° shooting angle, it has a nearly 100% detection rate and over 97% demodulation accuracy.

TABLE II: The reduced BER and the extraction time cost of different demodulation settings.

Number of subunits	1×1	2×2	3×3	4×4
Reduced BER (%)	0	3.891	4.182	4.045
Demodulation time (s)	1.48	3.13	5.52	8.64

5) *Extraction Time and Performance*: With the unoptimized Matlab implementation on a PC with AMD Ryzen 5 1600X CPU and NVIDIA GTX 1050 GPU, the average OAcodes extraction time with fast demodulation is 1.48s. For the enhanced demodulation, as shown in Table II, the number of subunits affects the average extraction time and demodulation accuracy. The enhanced demodulation with 1×1 subunit is equivalent to the fast demodulation, having the same time cost and performance, reducing 0% bit error rate (BER). With the increase in the number of subunits, the enhanced demodulation reduces the BER by about 4% on average but takes more time. It should be noted that the increment of the reduced BER becomes slow when the number of subunits is over 2×2 . The possible reason is that 2×2 subunits recover almost all lens distortions in our camera-shooting conditions. As a result, it is appropriate to set the number of subunits of the enhanced demodulation as 2×2 .

TABLE III: The PSNR of comparative aesthetic 2D barcode schemes, one is calculated by their data area without position detection patterns and another one is calculated by the overall 2D barcode.

PSNR	PiCode	RA Code	QR images	OAcodes
Data area only	24.93	24.98	25.11	25.12
Overall	15.04	15.98	10.40	25.12

6) *Comparative Experiment with Aesthetic 2D Barcode*: We compare the proposed OAcodes with representative aesthetic 2D barcode schemes, including Picode [10], RA code [11], and QR image [9]. For all these methods, we use the same background images setup as Sec. VI-B1, including 100 images of different categories. For a fair comparison, the setting of comparative experiments refers to [10], [11]. Specifically, aesthetic 2D barcodes are captured at a fixed screen-camera distance of 8cm with two angles 0° or 20°. Their information capacity is set as 29×29 or 65×65 bits. Because the proposed OAcodes have no position detection patterns, as shown in Table III, we set the PSNR of their data area before and after beautification to the same level of 25.03 ± 0.1 . It should be noted that if position detection patterns are also considered as modifications to the background image, the PSNR of overall aesthetic 2D barcodes would decrease sharply except for the proposed OAcodes.

As shown in Table IV, in all settings, OAcodes have the highest extraction accuracy, which is higher than 99%. Moreover, the experiment result further proves the effectiveness of enhanced demodulation, especially in high information capacity. In the information capacity of 29×29 bits, compared to the fast demodulation, the enhanced demodulation improves the

TABLE IV: The extraction accuracy of OAcodes and comparative aesthetic 2D barcodes under different capacities and angle settings.

Accuracy (%)	PiCode	RA Code	QR image	OAcodes	
				Fast	Enhanced
29 × 29, 0°	88.361	97.366	95.321	97.079	99.516
29 × 29, 20°	89.096	96.328	94.532	98.472	99.409
65 × 65, 0°	94.998	97.686	90.793	91.667	99.567
65 × 65, 20°	87.051	96.712	80.430	89.258	97.887

TABLE V: The extraction accuracy of OAcodes and comparative deep data hiding methods under different angle settings.

Methods	Hidden Code	StegaStamp	OAcodes
PSNR	34.035	34.011	34.181
Acc (%), 0°	90.327	92.480	93.245
Acc (%), 20°	87.122	90.770	91.204

demodulation accuracy by around 2%. Yet in the capacity of 65×65 bits, the accuracy improved by enhanced demodulation is nearly 10%.

7) *Comparative Experiment with Deep Data Hiding*: In state-of-art deep data hiding methods, some of them have similar application scenarios to 2D barcodes. Thus, we select two representative methods with available evaluation models, including StegaStamp [44] and Hidden Code [47] to compare with OAcodes. Since we only have the pre-trained evaluation model of Hidden Code, whose parameters are hard to modify, we adjust the OAcodes and StegaStamp to align with Hidden Code for fair comparative experiments. Specifically, using these methods, we embed 196 bits in background images of 256 × 256 pixels and set their PSNR to the same level of 34.1±0.1. The image dataset and capturing setting are the same as Sec. VI-B6. As shown in Table V, under the two common camera angles, OAcodes has the highest extraction accuracy, higher than 90%. One possible reason is that OAcodes deals with lens distortion better with the enhanced demodulation. Additionally, we notice that Hidden Code and StegaStamp have relatively low extraction accuracy on logo images, probably because they are originally designed for natural images.

8) *Performance on Other Devices or Parameters*: One of the benefits of using an experimental setup with explicitly defined parameters is that for new devices or parameters, we can roughly estimate the performance of OAcodes. Specifically, the camera-shooting process can be modeled as Fig. 12, where an OAcodes with physical side length L_O is captured by the camera at distance d , and converted into an image with side length L_S on the camera sensor with resolution r . In this model, the final captured pixel size L_p of OAcodes, which considerably affects the OAcodes extraction result, has the following relationship with the above parameters roughly:

$$L_p \propto L_S \cdot r \propto \frac{L_O \cdot r}{d}, \quad (23)$$

which could be used to estimate the performance of OAcodes under the new parameters. For example, according to the experimental result in Sec. VI-B2, OAcodes with side length

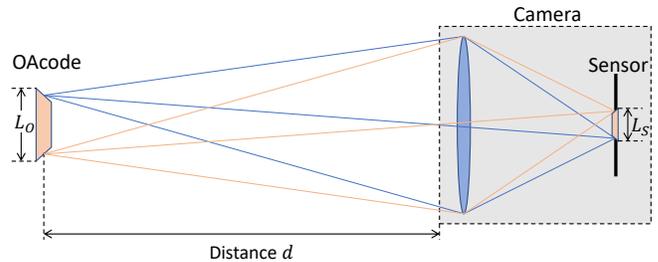


Fig. 12. Schematic diagram of the camera-shooting process of OAcodes.

$L_O = 5\text{cm}$ has 100% detection rate and 96.76% demodulation accuracy under the camera-shooting environment with distance $d = 15\text{cm}$ and camera resolution $r = 1280 \times 720$ pixels. According to Eq. (23), when OAcodes is displayed with double size $L_O = 10\text{cm}$ and the camera resolution is about three times larger 720p, OAcodes is supposed to have similar extraction performance at $2 \times 3 = 6$ times the distance, i.e., $15 \times 6 = 90\text{cm}$. We use different brands of phones with default camera resolutions to verify this assumption. As shown in Table VI, OAcodes has a longer available distance under better devices or parameters and roughly conforms to the relationship in Eq. (23).

TABLE VI: The detection rate and demodulation accuracy of OAcodes captured by different phones at a distance of 90cm.

Phones	Xiaomi Mi6	Redmi K40	Huawei P30 Pro	iPhone 11
Resolution	3016 × 4032	3000 × 3000	2976 × 2976	4032 × 2268
Det rate (%)	80	100	100	100
Acc (%)	89.355	94.785	93.681	97.100

VII. CONCLUSION

In this paper, we propose a novel overall aesthetic 2D barcode scheme, named OAcodes, which has no position detection patterns but achieves all functions that a complete aesthetic 2D barcode should have. Experimental results illustrate that, at certain screen-camera distances and angles, OAcodes could be quickly detected and demodulated. Additionally, the performance of OAcodes could be further improved with better devices and parameters. In our future work, we will explore the application of OAcodes in document printing and investigate a crop-resistant 2D barcode based on the framework of OAcodes.

REFERENCES

- [1] E. Ohbuchi, H. Hanaizumi, and L. A. Hock, "Barcode readers using the camera device in mobile phones," in *2004 International Conference on Cyberworlds*, 2004, pp. 260–265.
- [2] M. S. Hossain, X. Zhou, and M. F. Rahman, "Examining the impact of qr codes on purchase intention and customer satisfaction on the basis of perceived flow," *International Journal of Engineering Business Management*, vol. 10, p. 1847979018812323, 2018. [Online]. Available: <https://doi.org/10.1177/1847979018812323>
- [3] T. Cata, P. S. Patel, and T. Sakaguchi, "Qr code: A new opportunity for effective mobile marketing," *Journal of Mobile technologies, knowledge and society*, vol. 2013, p. 1, 2013.
- [4] S. Okazaki, H. Li, and M. Hirose, "Benchmarking the use of qr code in mobile promotion: three studies in japan," *Journal of Advertising Research*, vol. 52, no. 1, pp. 102–117, 2012.

- [5] "Research on qr code marketing: Effectiveness, adoption, and use cases," 2022. [Online]. Available: <https://www.qrcodechimp.com/qr-code-marketing-research/>
- [6] X. Min, K. Gu, G. Zhai, X. Yang, W. Zhang, P. Le Callet, and C. W. Chen, "Screen content quality assessment: overview, benchmark, and beyond," *ACM Computing Surveys (CSUR)*, vol. 54, no. 9, pp. 1–36, 2021.
- [7] K. Ma, S. Das, Z. Shu, and D. Samaras, "Learning from documents in the wild to improve document unwarping," in *ACM SIGGRAPH 2022 Conference Proceedings*, 2022, pp. 1–9.
- [8] A. J. Sellen and R. H. Harper, *The myth of the paperless office*. MIT press, 2003.
- [9] G. J. Garateguy, G. R. Arce, D. L. Lau, and O. P. Villarreal, "Qr images: Optimized image embedding in qr codes," *IEEE Transactions on Image Processing*, vol. 23, no. 7, pp. 2842–2853, 2014.
- [10] W. Huang and W. H. Mow, "Picocode: 2d barcode with embedded picture and vicocode: 3d barcode with embedded video," in *Proceedings of the 19th annual international conference on Mobile computing & networking*, 2013, pp. 139–142.
- [11] C. Chen, B. Zhou, and W. H. Mow, "Ra code: A robust and aesthetic code for resolution-constrained applications," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 28, no. 11, pp. 3300–3312, 2018.
- [12] C. Chen, W. Huang, L. Zhang, and W. H. Mow, "Robust and unobtrusive display-to-camera communications via blue channel embedding," *IEEE Transactions on Image Processing*, vol. 28, no. 1, pp. 156–169, 2019.
- [13] "Qrcode," 2020. [Online]. Available: <https://www.qrcode.com/zh/index.html>
- [14] "Data matrix," 2020. [Online]. Available: https://en.wikipedia.org/wiki/Data_Matrix
- [15] G. Zhai and X. Min, "Perceptual image quality assessment: a survey," *Science China Information Sciences*, vol. 63, no. 11, pp. 1–52, 2020.
- [16] X. Min, K. Gu, G. Zhai, J. Liu, X. Yang, and C. W. Chen, "Blind quality assessment based on pseudo-reference image," *IEEE Transactions on Multimedia*, vol. 20, no. 8, pp. 2049–2062, 2017.
- [17] X. Min, G. Zhai, K. Gu, Y. Liu, and X. Yang, "Blind image quality estimation via distortion aggravation," *IEEE Transactions on Broadcasting*, vol. 64, no. 2, pp. 508–517, 2018.
- [18] X. Min, G. Zhai, J. Zhou, M. C. Farias, and A. C. Bovik, "Study of subjective and objective quality assessment of audio-visual signals," *IEEE Transactions on Image Processing*, vol. 29, pp. 6054–6068, 2020.
- [19] X. Min, K. Ma, K. Gu, G. Zhai, Z. Wang, and W. Lin, "Unified blind quality assessment of compressed natural, graphic, and screen content images," *IEEE Transactions on Image Processing*, vol. 26, no. 11, pp. 5462–5474, 2017.
- [20] X. Min, K. Gu, G. Zhai, M. Hu, and X. Yang, "Saliency-induced reduced-reference quality index for natural scene and screen content images," *Signal Processing*, vol. 145, pp. 127–136, 2018.
- [21] R. Cox, "Qartcodes," April 2012. [Online]. Available: <http://research.swtch.com/qart>
- [22] 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, pp. 1–8, 2013.
- [23] S. Lin, M. Hu, C. Lee, and T. Lee, "Efficient qr code beautification with high quality visual content," *IEEE Transactions on Multimedia*, vol. 17, no. 9, pp. 1515–1524, 2015.
- [24] N. Aliva, U. Peled, and F. Itamar, "Visualead," 2020. [Online]. Available: <https://www.visualead.com/>
- [25] Z. Baharav and R. Kakarala, "Visually significant qr codes: Image blending and statistical analysis," in *2013 IEEE International Conference on Multimedia and Expo (ICME)*, 2013, pp. 1–6.
- [26] Y. Lin, Y. Chang, and J. Wu, "Appearance-based qr code beautifier," *IEEE Transactions on Multimedia*, vol. 15, no. 8, pp. 2198–2207, 2013.
- [27] M. Xu, H. Su, Y. Li, X. Li, J. Liao, J. Niu, P. Lv, and B. Zhou, "Stylized aesthetic qr code," *IEEE Transactions on Multimedia*, vol. 21, no. 8, pp. 1960–1970, 2019.
- [28] 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*, 2016, pp. 904–915.
- [29] J.-C. Liu and H.-A. Shieh, "Toward a two-dimensional barcode with visual information using perceptual shaping watermarking in mobile applications," *Optical Engineering*, vol. 50, no. 1, p. 017002, 2011.
- [30] J. Jia, G. Zhai, J. Zhang, Z. Gao, Z. Zhu, X. Min, X. Yang, and G. Guo, "Embdn: An efficient multiclass barcode detection network for complicated environments," *IEEE Internet of Things Journal*, vol. 6, no. 6, pp. 9919–9933, 2019.
- [31] J. Jia, G. Zhai, P. Ren, J. Zhang, Z. Gao, X. Min, and X. Yang, "Tiny-bdn: An efficient and compact barcode detection network," *IEEE Journal of Selected Topics in Signal Processing*, vol. 14, no. 4, pp. 688–699, 2020.
- [32] J. Zhang, X. Min, J. Jia, Z. Zhu, J. Wang, and G. Zhai, "Fine localization and distortion resistant detection of multi-class barcode in complex environments," *Multim. Tools Appl.*, vol. 80, no. 11, pp. 16 153–16 172, 2021. [Online]. Available: <https://doi.org/10.1007/s11042-019-08578-x>
- [33] D. W. Inc. and Elkhart, "Securing margin," 2020. [Online]. Available: <https://www.qrcode.com/en/howto/code.html>
- [34] S. Kang, S. D. Kim, and M. Kim, "Structural-information-based robust corner point extraction for camera calibration under lens distortions and compression artifacts," *IEEE Access*, vol. 9, pp. 151 037–151 048, 2021.
- [35] F. A. Petitcolas, R. J. Anderson, and M. G. Kuhn, "Attacks on copyright marking systems," in *International workshop on information hiding*. Springer, 1998, pp. 218–238.
- [36] F. A. Petitcolas, "Watermarking schemes evaluation," *IEEE Signal Process. Mag.*, vol. 17, no. 5, pp. 58–64, 2000.
- [37] F. A. P. Petitcolas, R. J. Anderson, and M. G. Kuhn, "Information hiding—a survey," *Proceedings of the IEEE*, vol. 87, no. 7, pp. 1062–1078, 1999.
- [38] C. I. Podilchuk and E. J. Delp, "Digital watermarking: algorithms and applications," *IEEE Signal Process. Mag.*, vol. 18, no. 4, pp. 33–46, 2001.
- [39] H. Fang, W. Zhang, Z. Ma, H. Zhou, S. Sun, H. Cui, and N. Yu, "A camera shooting resilient watermarking scheme for underpainting documents," *IEEE Trans. Circuits Syst. Video Technol.*, 2019.
- [40] H. Fang, W. Zhang, H. Zhou, H. Cui, and N. Yu, "Screen-shooting resilient watermarking," *IEEE Trans. Inf. Forensics Security*, vol. 14, no. 6, pp. 1403–1418, 2018.
- [41] A. Cheddad, J. Condell, K. Curran, and P. Mc Kevitt, "Digital image steganography: Survey and analysis of current methods," *Signal Processing*, vol. 90, no. 3, pp. 727–752, 2010. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0165168409003648>
- [42] N. Provos and P. Honeyman, "Hide and seek: an introduction to steganography," *IEEE Security & Privacy*, vol. 1, no. 3, pp. 32–44, 2003.
- [43] N. Subramanian, O. Elharrouss, S. Al-Maadeed, and A. Bouridane, "Image steganography: A review of the recent advances," *IEEE Access*, vol. 9, pp. 23 409–23 423, 2021.
- [44] M. Tancik, B. Mildenhall, and R. Ng, "Stegastamp: Invisible hyperlinks in physical photographs," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020.
- [45] H. Fang, D. Chen, F. Wang, Z. Ma, H. Liu, W. Zhou, W. Zhang, and N. Yu, "Tera: Screen-to-camera image code with transparency, efficiency, robustness and adaptability," *IEEE Transactions on Multimedia*, vol. 24, pp. 955–967, 2022.
- [46] J. Jia, Z. Gao, K. Chen, M. Hu, X. Min, G. Zhai, and X. Yang, "Rihoop: Robust invisible hyperlinks in offline and online photographs," *IEEE Transactions on Cybernetics*, vol. 52, no. 7, pp. 7094–7106, 2022.
- [47] J. Jia, Z. Gao, D. Zhu, X. Min, G. Zhai, and X. Yang, "Learning invisible markers for hidden codes in offline-to-online photography," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022, pp. 2273–2282.
- [48] Z. Ma, W. Zhang, H. Fang, X. Dong, L. Geng, and N. Yu, "Local geometric distortions resilient watermarking scheme based on symmetry," *IEEE Transactions on Circuits and Systems for Video Technology*, pp. 1–1, 2021.
- [49] R. Pickholtz, D. Schilling, and L. Milstein, "Theory of spread-spectrum communications - a tutorial," *IEEE Transactions on Communications*, vol. 30, no. 5, pp. 855–884, 1982.
- [50] T. Chapman, E. Larsson, P. von Wrycza, E. Dahlman, S. Parkvall, and J. Sköld, "Chapter 3 - cdma transmission principles," in *HSPA Evolution*. Oxford: Academic Press, 2015, pp. 35–48. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/B978080899969200003X>
- [51] J. S. Lim, *Two-Dimensional Signal and Image Processing*. Englewood Cliffs, NJ, Prentice Hall., 1990, p. 548, equations 9.26, 9.27, and 9.29.
- [52] "Belgalogos dataset," March 2014. [Online]. Available: <http://www-sop.inria.fr/members/Alexis.Joly/BelgaLogos/BelgaLogos.html>
- [53] "Cvg-ugr image database," March 2014. [Online]. Available: <http://decsai.ugr.es/cvg/dbimagenes/>