

Validation of optical codes based on 3D nanostructures

Artur Carnicer^{*a}, Bahram Javidi^b

^aUniversitat de Barcelona, Departament de Física Aplicada, Martí i Franquès 1, 08028 Barcelona, Catalunya (Spain); ^bUniversity of Connecticut, Electrical & Computer Engineering Dept., 371 Fairfield Way, Storrs, CT 06269 (USA)

ABSTRACT

Image information encoding using random phase masks produce speckle-like noise distributions when the sample is propagated in the Fresnel domain. As a result, information cannot be accessed by simple visual inspection. Phase masks can be easily implemented in practice by attaching cello-tape to the plain-text message. Conventional 2D-phase masks can be generalized to 3D by combining glass and diffusers resulting in a more complex, physical unclonable function. In this communication, we model the behavior of a 3D phase mask using a simple approach: light is propagated through glass using the angular spectrum of plane waves whereas the diffuser is described as a random phase mask and a blurring effect on the amplitude of the propagated wave. Using different designs for the 3D phase mask and multiple samples, we demonstrate that classification is possible using the k-nearest neighbors and random forests machine learning algorithms.

Keywords: Optical security, Speckle, Machine vision.

1. INTRODUCTION

In the last 20 years, a large number of light-based techniques for security and encryption problems have been suggested [1-4]. Very recently, nanostructured devices have been proposed for producing physical unclonable functions. Among other techniques, the information is encoded by using metal nanoparticles or thin layer structures. In practical applications, random phase masks attached to the image to be secured make the information inaccessible once the light interacting with the sample is propagated. These devices display a distinctive behavior under polarized light and consequently, the use of illuminating sources with different polarizations increase the degrees of freedom of the cryptosystem resulting in a gadget which is more complex and difficult to attack [5-8].

3D optical random phase masks were proposed some years ago [9]. These ensembles are produced by combining glass and diffuser materials. They are simple to produce and very difficult to copy. As in nanoparticle-based devices, the coded information is not accessible once the light interacting with the sample is propagated. The irradiance of the propagated signal is usually a speckle-like noise distribution and consequently, statistical analyses of the recorded light can be performed. Note that small differences in the noise pattern enable to distinguish the same encoded message but attached to a different random code. For this kind of problems, classification methods based on machine learning techniques can be used [10]. In this way, it is possible to discriminate among messages with the correct encoding (true class) and counterfeit information (false class).

In this communication we present a simplified numerical model that approximately describes the behavior of 3D optical random phase codes. In particular, we show that encoded messages can be distinguished using machine learning algorithms under certain propagation conditions.

2. A NUMERICAL MODEL FOR 3D RANDOM PHASE CODES

In this section we present a computer simulation to model the behavior of 3D random phase masks [9]. This calculation provides more insight on the results presented in reference [11]. In that paper, we used 3D phase masks to encode information in such a way that it cannot be read after propagation. Despite the speckle noise embeds and hides the original data, these signals are statistically labelled so they can be successfully classified using machine learning algorithms. In [11], Random Forests Classification [12] was used for discriminating among samples classes using experimentally obtained data.

* artur.carnicer@ub.edu

In order to produce a correct model of such devices, light propagation through a diffuser should be taken into account. However, the analysis of the transmitted light through an inhomogeneous system is very complex. Here, we propose a simplified model based on combining phase masks, glass slides and blurring, that can be considered as a first-order approximation.

Fig. 1(a) depicts the three 3D random codes designs used as numerical tests. The image used (a QR code) is shown in Fig. 1(b). Code 1 is generated by attaching a phase-only random mask ϕ_1 to $Q(x,y)$. For convenience, the complementary image of the QR was used. Code 2 was simulated by combining random phase masks ϕ_2 and ϕ_3 and glass: light is propagated through a 2 mm-wide microscope slide with index of refraction $n=1.5195$, corresponding to a BK7 glass illuminated at $\lambda=532$ nm. Code 3 is calculated in an equivalent way according to the corresponding diagram in Fig. 1(a). According to the Fresnel Laws, the amplitude transmission of the ensemble is

$$t = \frac{4n_{\text{air}}n_{\text{BK7}}}{(n_{\text{air}} + n_{\text{BK7}})^2}, \quad (1)$$

and therefore, the transmittance for codes 2 and 3 is $t_2=0.9575$ and $t_3=0.9168$ respectively. Blurring produced by the diffuser is simulated by convolving the modulus of the propagated beam by an $N \times N$ constant kernel. Propagation is calculated by means of the scalar plane wave power spectrum formula [13],

$$E(x, y, z) = \text{FT}^{-1} \left[\text{FT} [E(x, y, 0)] H(u, v, z) \right]$$

$$H(u, v, z) = \exp \left(i \frac{2\pi}{\lambda} z \sqrt{1 - \lambda^2 u^2 - \lambda^2 v^2} \right). \quad (2)$$

where FT stands for the Fourier transform operator, u and v are the spatial frequencies, $E(x,y,0)$ is the scalar electric field at the reference plane, $E(x,y,z)$ is the propagated beam and $H(u,v,z)$ is the transference function. Note that the wavelength is λ/n when the beam is propagated in a medium with refractive index n . Table 1 displays a summary of variables and values used in the simulation.

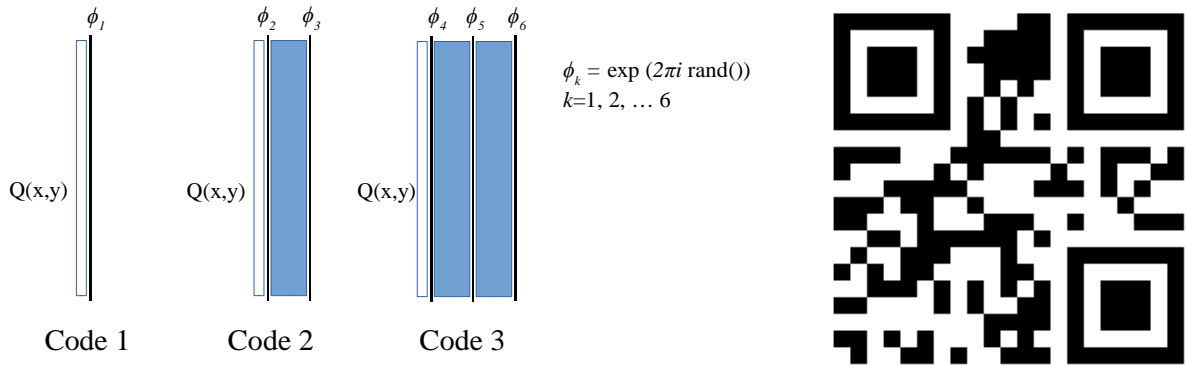


Figure 1. (a) Sketch of the three 3D codes masks used in the simulation; (b) test image: the QR encodes the message 0034934021143.

Table 1: Propagation variables

Wavelength λ	520 nm
Size of the code	15 x 15 mm ²
refractive index n_{BK7}	1.5195
Slide width	2 mm
Propagated distance z	250, 400 mm
Number of pixels	290 x 290 pixels
Blurring kernel	21 x 21 pixels
Camera depth	256 gray levels
Number of generated codes per class	30

Calculations are performed as follows: light propagates through one of the three encoding devices [Fig. 1(a)]. Then, the beams are propagated a distance $z=250$ mm and the light irradiance is recorded by means of an 8-bit camera. The generated distributions are shown in Figs 2(a-c). Note that the irradiance displays a speckle noise pattern and the information encoded in the QR is no accessible by simple visual inspection [see Fig2. 2(d-f)]. Figures 2(g-i) show the amplitude histograms (the square root of the irradiance). In order to provide enough cases, 30 different codes for each class have been calculated. Note that histograms present subtle differences that can be used to distinguish among the three classes.

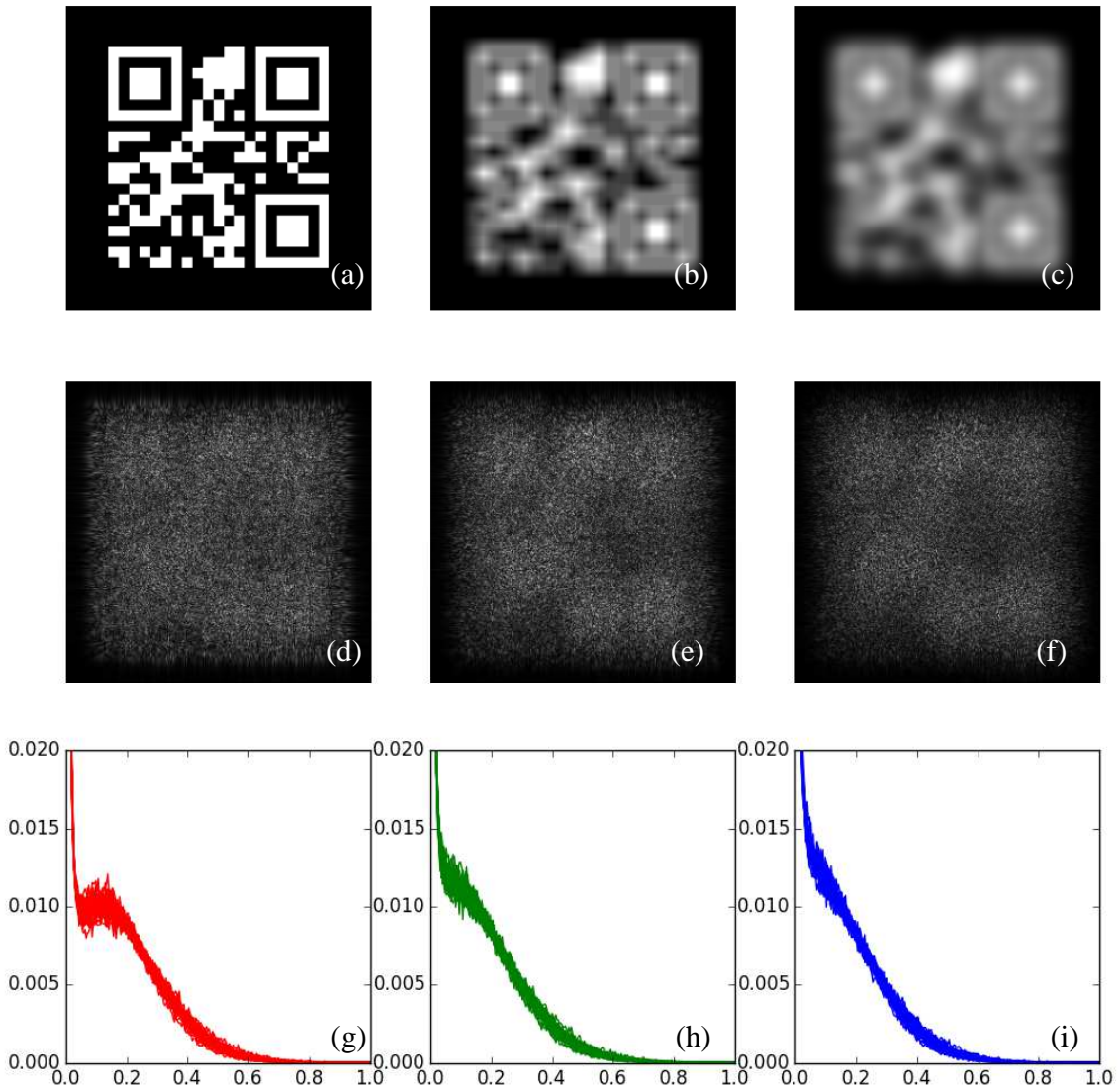


Figure 2. First row: Encoded light distributions trough (a) Code 1, (b) Code 2, (c) Code 3. Second row: Propagated light distributions at $z = 250$ mm (d) Code 1, (e) Code 2, (f) Code 3. Third row: Histograms of the propagated amplitudes: (g) Code 1, (h) Code 2, (i) Code 3.

Using a different propagation distance ($z=400$ mm), a second set of calculations have been carried out. The recorded distributions and the corresponding histograms are shown in Figs. 3(a-c) and 3(d-f), respectively. Note that in this case, the profile of the histograms clearly follows a Gamma probability distribution [14-16], i.e.

$$p(I) = \left(\frac{n_0}{\langle I \rangle} \right)^{n_0} \frac{I^{n_0-1} \exp(-I n_0 / \langle I \rangle)}{\Gamma(n_0)} \quad (3)$$

where $I, \langle I \rangle$ and σ are the intensity data points, its average and the corresponding standard deviation respectively; Parameter n_0 is defined as $n_0 = (\langle I \rangle / \sigma)^2$ and $\Gamma()$ stands for the Gamma special function.

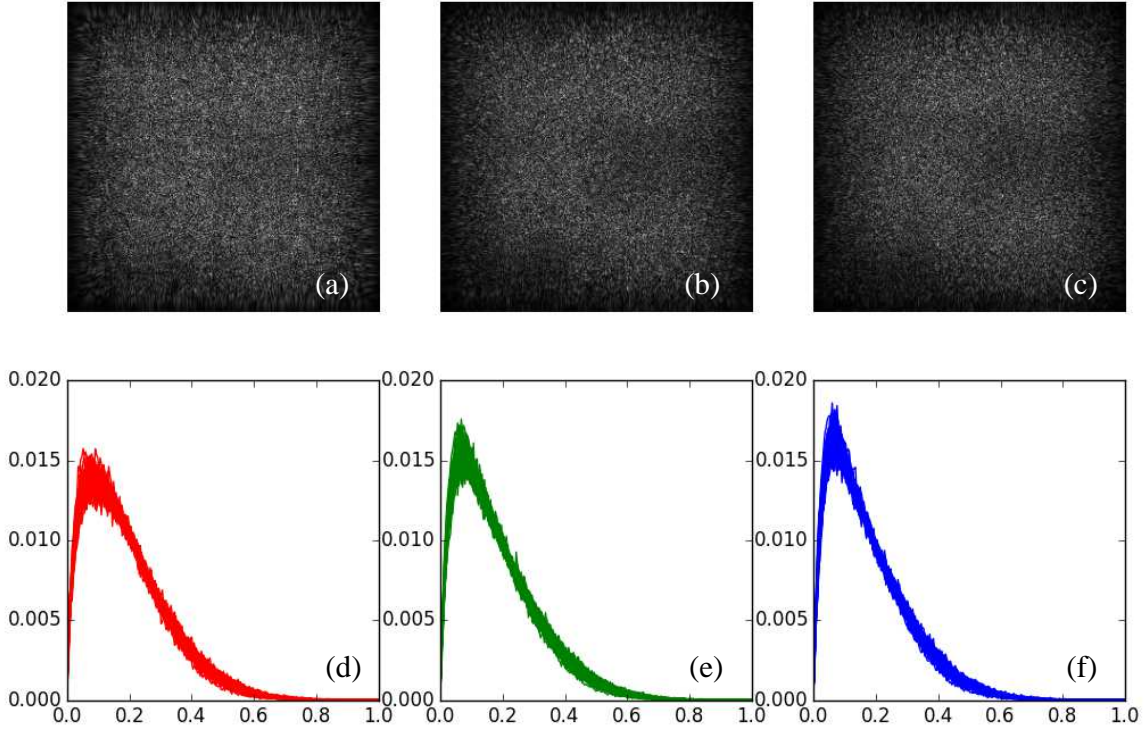


Figure 3. First row: Propagated light distributions $z = 400$ mm (a) Code 1, (b) Code 2, (c) Code 3. Second row: Histograms of the propagated amplitudes (d) Code 1, (e) Code 2, (f) Code 3.

3. CLASSIFICATION

Using the information provided by the histograms, it is possible to distinguish among the three different classes of codes. As explained above, 30 different codes for each class (Code 1, 2, and 3) have been produced. After propagation, the corresponding amplitude histograms are used to feed a machine learning classification system. In the present work two algorithms have been considered: k-Nearest Neighbors (kNN) with $k=1$ [17] and Random Forest (RF) with 100 trees [12]. Classification is performed by splitting the dataset into training and test subsets. Validation is carried out 10 times. In each trial, the samples used in the training or test sets are randomly selected. Classification accuracy is calculated by means of the ratio of the number of correctly classified samples, divided by the total number of test samples. Calculations were carried out using the scikit-learn library [18].

Table 2 summarizes the results obtained: for $z = 250$ mm, both kNN and RF produce very good results. In particular kNN is able to produce an error-free classification. Nevertheless, classification accuracy is very low for $z = 400$ mm. An alternative approach for classification has been considered. Instead of using the 256 values of the histogram, several statistical parameters have been obtained from the histogram distribution: mean, standard deviation, kurtosis and skewness. On top of that, the entropy of the irradiance has also been considered. These five statistics are used to train a system using kNN ($k=1$) and RF (100 trees). As in the previous case, each class is composed of 30 samples. Validation is performed 10 times using a hold-out strategy. In this case, RF provides almost error-free classification for $z = 250$ mm. When irradiance is recorded at $z = 400$ mm, both kNN and RF fails and a bad classification accuracy is obtained.

At first sight, it seems that histograms for the $z = 250$ mm case are somewhat different whereas the histograms generated for the $z = 400$ mm case are almost undistinguishable.

Table 2: Classification results

Data	Distance	Nearest neighbor	Random Forests
Histograms	250 mm	1.000	0.964
	400 mm	0.780	0.760
Statistics	250 mm	0.749	0.989
	400 mm	0.487	0.733
# members of each class			30
# trials for validation			10

4. CONCLUDING REMARKS

In this communication, we proposed a simplified numerical model for understanding the behavior of optical 3D codes based on diffusers. These structures encode the image in a unique way and after propagation, the irradiance presents a speckle-like noise pattern. As a result, the original information is no longer accessible. The statistics of these distributions is analyzed and the data provided is used to feed machine learning classification methods. We demonstrated that using kNN and RF is possible to distinguish among the signals produced by different classes of codes. We found that classification is possible for a certain range of distances.

ACKNOWLEDGMENTS

AC acknowledges financial support from Ministerio de Economía y Competitividad (MINECO), projects FIS2013-46475-C3-2-P and FIS2016-75147-C3-1-P.

REFERENCES

- [1] O. Matoba, T. Nomura, E. Perez-Cabre, M. S. Millan and B. Javidi, "Optical techniques for information security," Proc. IEEE **97**, 1128-1148 (2009).
- [2] A. Alfalou and C. Brosseau, "Optical image compression and encryption methods," Adv. Opt. Photonics **1**, 589-636 (2009)
- [3] W. Chen, B. Javidi and X. Chen, "Advances in optical security systems," Adv. Opt. Photonics **6**, 120-155 (2014).
- [4] B. Javidi, A. Carnicer, M. Yamaguchi, T. Nomura, E. Pérez-Cabré, M. S. Millán, N. K. Nishchal, R. Torroba, J. F. Barrera, W. He, X. Peng, A. Stern, Y. Rivenson, A. Alfalou, C. Brosseau, C. Guo, J. T. Sheridan, G. Situ, M. Naruse, T. Matsumoto, I. Juvells, E. Tajahuerce, J. Lancis, W. Chen, X. Chen, P. W. H. Pinkse, A. P. Mosk, A. Markman "Roadmap on Optical Security," J. Opt. **16**, 083001 (2016).
- [5] A. Markman, B. Javidi and M. Tehranipoor, "Photon-counting security tagging and verification using optically encoded QR codes," IEEE Photonics J. **6**, 1-9 (2014).
- [6] A. Carnicer, A. Hassanfiroozi, P. Latorre-Carmona, Y.-P. Huang and B. Javidi, "Security authentication using phase-encoded nanoparticle structures and polarized light," Opt. Lett. **40**, 2135-138 (2015).
- [7] A. Carnicer, O. Arteaga, E. Pascual, A. Canillas, S. Vallmitjana, B. Javidi and E. Bertran, "Optical security verification by synthesizing thin films with unique polarimetric signatures," Opt. Lett. **40**, 5399-5402, (2015).
- [8] A. Carnicer, O. Arteaga, J. M. Suñé and B. Javidi, "Authentication of gold nanoparticle encoded pharmaceutical tablets using polarimetric signatures," Opt. Lett. **41**, 4507-4510 (2016).
- [9] O. Matoba, S. Matsuki and K. Nitta, "Secure data storage by three-dimensional absorbers in highly scattering volume medium," in Journal of Physics: Conference Series **139**, 1 (2008).
- [10] N. Japkowicz and M. Shah, *Evaluating learning algorithms: a classification perspective* (Cambridge University, 2011).
- [11] A. Markman, A. Carnicer and B. Javidi, "Security authentication with a three-dimensional optical phase code using random forest classifier," J. Opt. Soc. Am. A **33**, 1160-1165 (2016)
- [12] L. Breiman, "Random forests," Machine learning **45**, 5-32 (2001).
- [13] J. W. Goodman, *Introduction to Fourier optics* (Roberts and Company, 2005).
- [14] J. C. Dainty, "The statistics of speckle patterns," Progress in Optics **14**, 1-46 (1977).

- [15] J. W. Goodman, *Statistical optics* (John Wiley & Sons, 2015).
- [16] A. Carnicer, B. Javidi, "Optical Security and Authentication using Nanoscale and Thin Film structures," *Adv. Opt. Photonics* **9**(1) (2017).
- [17] T. Cover and P. Hart, "Nearest neighbor pattern classification," *IEEE Trans. Inf. Theory* **13**, 21-27 (1967).
- [18] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, É. Duchesnay, "Scikit-learn: Machine learning in Python," *J. Mach. Learn. Res* **12**, 2825-2830 (2011).