Resistance of the double random phase encryption against various attacks

Yann Frauel[1*], **Albertina Castro** [2], **Thomas J. Naughton**[3], **and Bahram Javidi**[4]

[1] *Departamento de Ciencias de la Computación*
Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas
Universidad Nacional Autónoma de México
México, DF 04510, Mexico.
[2] *Instituto Nacional de Astrofísica, Óptica y Electrónica*
Apdo. Postal 51, Puebla, Pue. 72000, Mexico
[3]*Dept. of Computer Science, National University of Ireland, Maynooth*
County Kildare, Ireland
University of Oulu, RFMedia Laboratory, Oulu Southern Institute
Vierimaantie 5, 84100 Ylivieska, Finland.
[4] *Electrical & Computer Engineering Dept., University of Connecticut*
371 Fairfield Road, Unit 2157, Storrs CT 06269-2157, USA.
** Corresponding author: yann@leibniz.iimas.unam.mx*

Abstract: Several attacks are proposed against the double random phase encryption scheme. These attacks are demonstrated on computer-generated ciphered images. The scheme is shown to be resistant against brute force attacks but susceptible to chosen and known plaintext attacks. In particular, we describe a technique to recover the exact keys with only two known plain images. We compare this technique to other attacks proposed in the literature.

OCIS codes: (070.2580) Fourier optics; (070.4560) Optical data processing; (200.4740) Optical processing; (999.9999) Optical encryption.

References and links

1. Ph. Réfrégier and B. Javidi, "Optical image encryption based on input plane and Fourier plane random encoding," Opt. Lett. **20,** 767-769 (1995).
2. B. Javidi and E. Ahouzi, "Optical security system with Fourier plane encoding," Appl. Opt. **37,** 6247-6255 (1998).
3. J.-W. Han, C.-S. Park, D.-H. Ryu, and E.-S. Kim, "Optical image encryption based on XOR operations," Opt. Eng. **38,** 47-54 (1999).
4. J. L. Horner and B. Javidi, Opt. Eng. **38,** Special issue on Optical security, 1999.
5. S. Lai and M. A. Neifeld, "Digital wavefront reconstruction and its application to image encryption," Opt. Commun. **178,** 283-289 (2000).
6. Y. Li, K. Kreske, and J. Rosen, "Security and encryption optical systems based on a correlator with significant output images," Appl. Opt. **39,** 5295-5301 (2000).
7. X. Peng, Z. Cui and T. Tan, "Information encryption with virtual-optics imaging system," Opt. Commun. **212,** 235-245 (2002).
8. B. M. Hennelly and J. T. Sheridan, "Image encryption and the fractional Fourier transform," Optik **114,** 251-265 (2003).
9. O. Matoba and B. Javidi, "Secure three-dimensional data transmission and display," Appl. Opt. **43,** 2285-2291 (2004).
10. B. Javidi, ed., *Optical and Digital Techniques for Information Security* (Springer Verlag, New York, 2005).
11. L. G. Neto, and Y. Sheng, "Optical implementation of image encryption using random phase encoding," Opt. Eng. **35,** 2459-2463 (1996).

#80383 - $15.00 USD Received 23 Feb 2007; revised 27 Apr 2007; accepted 10 May 2007; published 30 Jul 2007

12. N. Towghi, B. Javidi, and Z. Luo, "Fully phase encrypted image processor," J. Opt. Soc. Am. A **16,** 1915-1927 (1999).
13. O. Matoba and B. Javidi, "Encrypted optical storage with wavelength-key and random phase codes," Appl. Opt. **38,** 6785-6790 (1999).
14. O. Matoba and B. Javidi, "Encrypted optical memory system using three-dimensional keys in the Fresnel domain," Opt. Lett. **24,** 762-764 (1999).
15. E. Tajahuerce and B. Javidi, "Encrypting three-dimensional information with digital holography," Appl. Opt. **39,** 6595-6601 (2000).
16. G. Unnikrishnan, J. Joseph, and K. Singh, "Optical encryption by double-random phase encoding in the fractional Fourier domain," Opt. Lett. **25**, 887–889 (2000).
17. G. Unnikrishnan and K. Singh, "Optical encryption using quadratic phase systems," Opt. Commun. **193,** 51-67 (2001).
18. S. Liu, L. Yu, B. Zhu, "Optical image encryption by cascaded fractional Fourier transforms with random phase filtering, " Opt. Commun. **187** 57-63 (2001).
19. J. Ohtsubo and A. Fujimoto, "Practical image encryption and decryption by phase-coding technique for optical security systems," Appl. Opt. **41,** 4848-4855 (2002).
20. H. T. Chang, W. C. Lu, and C. L. Kuo, "Multiple-phase retrieval for optical security systems by use of random-phase encoding," Appl. Opt. **41,** 4825-4834 (2002).
21. B. Zhu, H. Zhao, and S. Liu, "Image encryption based on pure intensity random coding and digital holography technique," Optik **114,** 95-99 (2003).
22. G. Situ and J. Zhang, "Double random-phase encoding in the Fresnel domain," Opt. Lett. **29,** 1584-1586 (2004).
23. X. Wang, D. Zhao, L. Chen, "Image encryption based on extended fractional Fourier transform and digital holography technique," Opt. Commun. **260**, 449–453 (2006).
24. H. Suzuki, M. Yamaguchi, M. Yachida, N. Ohyama, H. Tashima, and T. Obi, "Experimental evaluation of fingerprint verification system based on double random phase encoding," Opt. Express **14,** 1755-1766 (2006).
25. A. Carnicer, M. Montes-Usategui, S. Arcos, and I. Juvells, "Vulnerability to chosen-cyphertext attacks of optical encryption schemes based on double random phase keys," Opt. Lett. **30,** 1644-1646 (2005).
26. Y. Frauel, A. Castro, T. J. Naughton, and B. Javidi, "Security analysis of optical encryption," in *Unmanned/Unattended Sensors and Sensor Networks II*, E. M. Carapezza, ed., Proc. SPIE **5986,** 25-34 (2005).
27. U. Gopinathan, D. S. Monaghan, T. J. Naughton, and J. T. Sheridan, "A known-plaintext heuristic attack on the Fourier plane encryption algorithm," Opt. Express **14,** 3181-3186 (2006).
28. X. Peng, P. Zhang, H. Wei, and B. Yu, "Known-plaintext attack on optical encryption based on double random phase keys," Opt. Lett. **31,** 1044-1046 (2006).
29. X. Peng, H. Wei, and P. Zhang, "Chosen-plaintext attack on lensless double-random phase encoding in the Fresnel domain," Opt. Lett. **31,** 3261-3263 (2006).
30. H. Beker and F. Piper, *Cipher systems* (Van Nostrand, London, 1982).
31. S. R. Blackburn, S. Murphy, and K. G. Paterson, "Comments on 'Theory and applications of cellular automata in cryptography'," IEEE Trans. Comp. **46,** 637-638 (1997).
32. W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical recipes in C* (Cambridge University Press, Cambridge, 1992), Chap. 2.

1. Introduction

Over the past few decades, the inherent parallelism of optics has prompted a great deal of research on optical information processing. Among all the possible applications, a very promising direction is optical encryption [1]-[24]. With their pioneering work on double random phase (DRP) encoding, Réfrégier and Javidi [1] paved the way for many following proposals of optical security and encryption systems. This very DRP scheme has spawned several variation techniques and has been applied in many situations [11]-[24]. However only recently has the security of DRP started to be thoroughly analyzed and a few weaknesses have started to appear [25]-[29]. Carnicer et al. presented a chosen ciphertext attack where an attacker can retrieve the key by inducing a legitimate user in deciphering many specially crafted ciphered images [25]. Gopinathan et al. [27] and Peng et al. [28] described two techniques to recover the key knowing a single plain image and the corresponding ciphered image. In the first case the key is obtained through a simulated annealing process, while in the second case a phase retrieval algorithm is used. A variant of the DRP encryption is also attacked in [29]. In this paper, we comment on several possible attacks on the DRP encryption scheme and in particular we present a known plaintext attack that requires two known plain images [26], but that is able to retrieve the exact

#80383 - $15.00 USD Received 23 Feb 2007; revised 27 Apr 2007; accepted 10 May 2007; published 30 Jul 2007

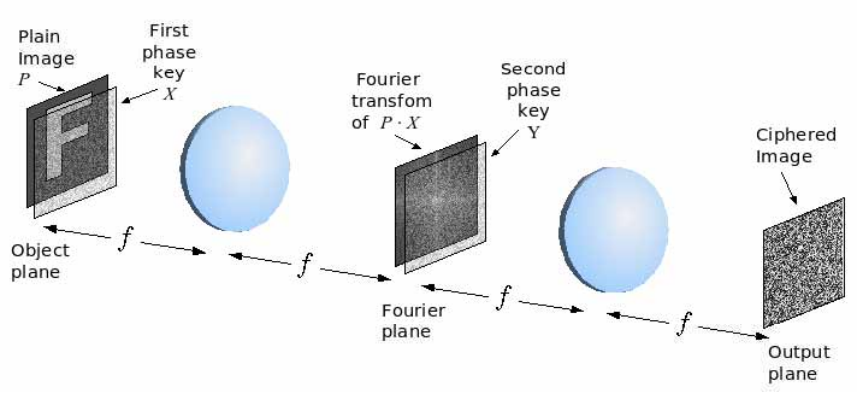

Fig. 1. Principle of the double random phase encryption scheme.

key rather than an approximation of it as in Refs. [27] and [28]. We demonstrate the feasibility of the proposed attacks using computer-generated images.

In section 2, we recall the principle of the double random phase encryption. In section 3 we consider brute force attacks to the system. In sections 4 and 5 we describe chosen plaintext attacks and known plaintext attacks respectively. Section 6 gives a discussion of the various attacks presented in this and other papers. Conclusions are presented in Section 7.

2. Overview of the double random phase encryption scheme

The original double random phase encryption technique was proposed in [1]. It is described in Fig. 1. The image to be encrypted, or plain image P, is immediately followed by a first random phase mask, which is the first key X. Both the image and the mask are located in the object focal plane of a first lens. In the image focal plane of this lens is therefore obtained the Fourier transform (FT) of the product $P \cdot X$. This product is then multiplied by another random phase mask that is the second key Y. Lastly, another FT is performed by a second lens to return to the spatial domain. Since the last FT does not add anything to the security of the system, we will perform all our analyses in the Fourier plane. The ciphered image C is then

$$C = Y \cdot \mathscr{F}(P \cdot X), \tag{1}$$

where $\mathscr{F}$ stands for the Fourier transform operation. In most of the paper, we will assume that P is a gray-level image — the so-called amplitude encoding. However, the last attacks we will present also work with the phase encoding variant [12] where the image is pure phase (see Section 6).

3. Brute force attacks

3.1. *Exact decryption*

The first obvious attack against any encryption system consists of trying every possible key until finding the correct one. Assuming that both phase masks have a size of $\sqrt{N} \times \sqrt{N}$ pixels and that each pixel has L possible phase values, the number of attempts required to retrieve both keys is of the order of L^{2N}. For any practical values of L and N, this number is huge and it increases exponentially with the number of pixels N which makes this attack computationally intractable. For instance, let us fix $L = 16$ phase levels and $N = 100 \times 100$ pixels. Then the number of keys to try would be $16^{20000} \sim 10^{24000}$, which is inconceivable. In practice, however, it is not always necessary to consider all of these combinations.

#80383 - $15.00 USD Received 23 Feb 2007; revised 27 Apr 2007; accepted 10 May 2007; published 30 Jul 2007

Thus, in the most common case of amplitude encoding [1], we do not need to know the first key X to decipher an image C. Indeed, if we know the second key Y, we can find

$$P \cdot X = \mathscr{F}^{-1}\left(C \div Y\right), \tag{2}$$

where $\mathscr{F}^{-1}$ denotes the inverse Fourier transform and $\div$ stands for the point-by-point division. Now because P is a pure magnitude function and X is a pure phase function, the plain image is easily recovered by

$$P = |P \cdot X| = \left|\mathscr{F}^{-1}\left(C \div Y\right)\right|, \tag{3}$$

where $|\,|$ is the modulus. This property leads to the first reduced brute-force attack:

Attack 1: Try every possible instance of the second key Y, while ignoring the first key X.

The number of trials required by this attack is L^N instead of L^{2N}. This represents a great simplification since the number of combinations is reduced by a factor L^N. In the case of the numerical example used above, the number of trials is $16^{10000} \sim 10^{12000}$.

3.2. Approximate decryption

In addition to attack 1, it is possible to look for an approximate version of the key Y instead of its exact value. The first possible simplification consists of reducing the number of phase levels, at the cost of a loss of quality in the decrypted image. For instance, we can look for keys with only two phase levels.

Attack 2: Try only binary phase keys for the second key Y; ignore the first key X.

The number of possible combinations is then 2^N instead of L^N. For the example of section 3.1, the number of combinations to try would be $2^{10000} \sim 10^{3000}$. Figure 2 shows the result of decrypting two 100×100 images with various reductions of phase levels of the key Y. The original key has 16 phase levels but the original image is recognizable even with a binarized phase key. However, the fewer phase levels, the more noise is introduced in the reconstruction.

Another possible simplification is to retrieve a partial window of the key Y instead of the full key [25, 26, 29]. Since the pixels outside the window are then ignored, deciphering with the partial window amounts to applying a low-pass filter to the deciphered image. The details of the original image are thus lost but large features can be recognized.

Attack 3: Brute force search of a partial window of the second key Y; ignore the first key X.

If the window has a total of $N_r < N$ pixels, then the number of combinations to try is L^{N_r} instead of L^N. Figure 3 presents decryptions of two 100×100 images using partial windows of the ciphered image with various sizes. In this case, retrieving a 30×30 window of the second key is sufficient to recognize the plain image, although the details are lost. The number of combinations is thus reduced to $16^{900} \sim 10^{1080}$.

Additionally, it is possible to combine Attacks 2 and 3 as follows:

Attack 4: Brute force search of a partial window of the second key Y trying a small number of phase levels; ignore the first key X.

For instance, if we use only three phase levels then the number of combinations to try is 3^{N_r} instead of the original L^N . Figure 4 presents decrypted images using ternary phase levels for

#80383 - $15.00 USD Received 23 Feb 2007; revised 27 Apr 2007; accepted 10 May 2007; published 30 Jul 2007

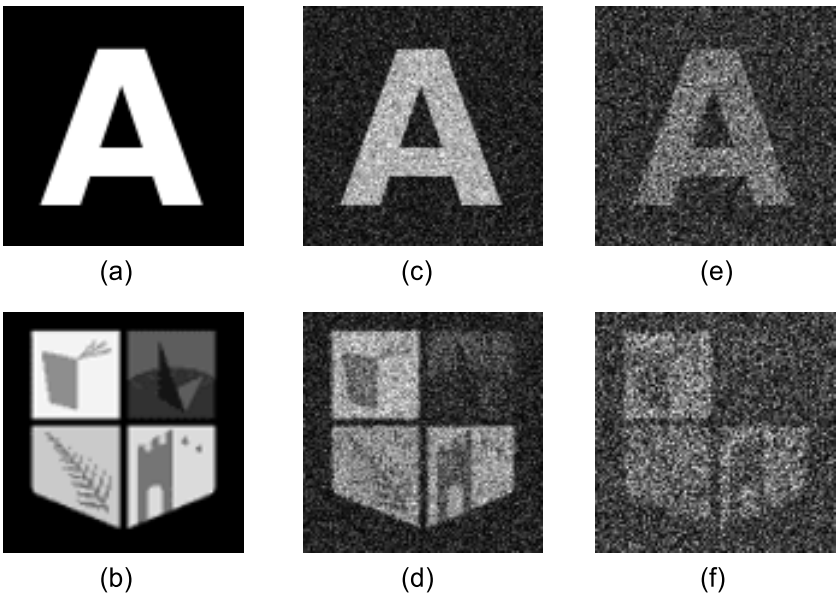

Fig. 2. Decryption of images encrypted with 16-level phase keys. (a)-(b) 16-level decryption key, (c)-(d) 4-level decryption key, and (e)-(f) 2-level decryption key.

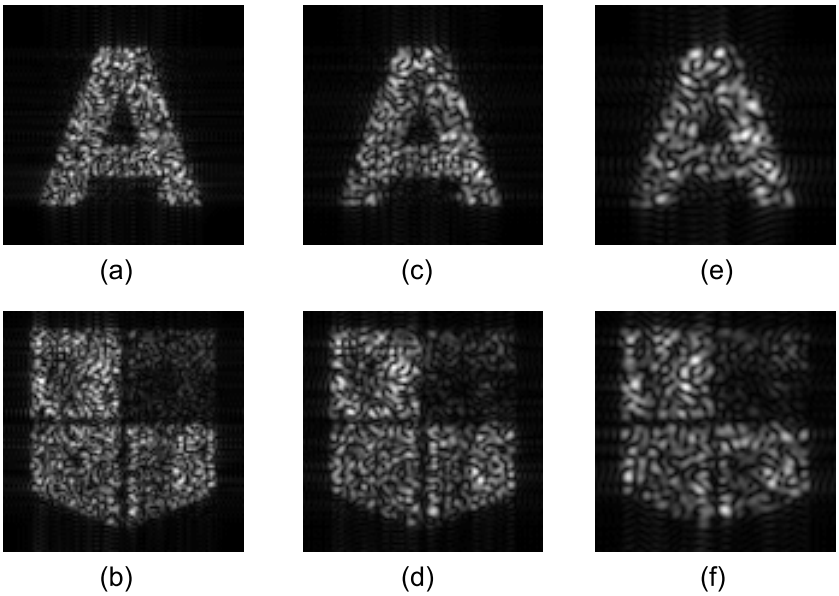

Fig. 3. Decryption using partial windows of the original 100 $\times$ 100 key. (a)-(b) 50 $\times$ 50 window, (c)-(d) 40 $\times$ 40 window, and (e)-(f) 30 $\times$ 30 window.

#80383 - $15.00 USD Received 23 Feb 2007; revised 27 Apr 2007; accepted 10 May 2007; published 30 Jul 2007

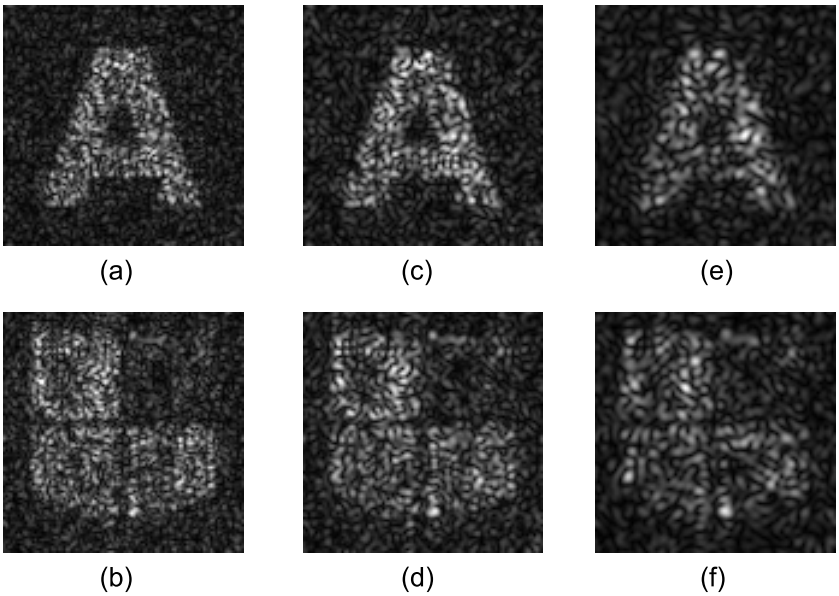

Fig. 4. Decryption using partial windows of the original 100 $\times$ 100 key and reduction to three phase levels. (a)-(b) 50 $\times$ 50 window, (c)-(d) 40 $\times$ 40 window, and (e)-(f) 30 $\times$ 30 window.

the second key and various window sizes. It can be seen that a gray-level image with details deteriorates faster than a binary image with a low-level of detail. In the case of Fig. 4(e), the number of trials would be $3^{900} \sim 10^{429}$, while the image is still recognizable. Although that number represents a dramatic reduction from the original 10^{24000} attempts, the number of trials remains huge. Therefore we can conclude that brute force attacks are intractable. However it should be kept in mind that the nominal security of the scheme is lower than one may naively expect.

4. Chosen plaintext attacks

We now assume that the attacker has the ability to trick a legitimate user of the system into encrypting particular images. In this case, as mentioned in [25] and [26], a very simple and yet very effective attack can be mounted:

Attack 5: Obtain the ciphered image corresponding to a Dirac delta function as plain image.

In other terms, the plain image is entirely black except for a single pixel. If the non-black pixel is centered, it is straightforward to see from Eq. (1) that the resulting ciphered image is — to a constant phase factor — the second key Y [26]. If the non-black pixel is not centered, then the ciphered image is the second key Y multiplied by a linearly varying phase. Since this linear phase is known (it only depends on the position of the non-black pixel), it is possible to extract the key Y. Thus, the system can be easily broken by encrypting a single chosen plain image. It can be argued that such a plain image can look suspicious to the authorized user that is to encrypt it. A variant can therefore be used:

Attack 6: Obtain and subtract the ciphered images corresponding to two plain images whose difference is a delta function.

#80383 - $15.00 USD Received 23 Feb 2007; revised 27 Apr 2007; accepted 10 May 2007; published 30 Jul 2007

This attack is made possible because the encryption scheme is linear. Indeed, If P_1, P_2 are two plain images with corresponding ciphered images C_1, C_2, and α, β are two scalar coefficients, it is easy to see from Eq. (1) that $\alpha P_1 + \beta P_2$ is encrypted to $\alpha C_1 + \beta C_2$. In the case of Attack 6, if $P_1 - P_2 = \delta$ then $C_1 - C_2$ is the ciphered image of a delta function — that is the second key Y — as in Attack 5. In order to disguise the attack even further, it is possible to use an arbitrary number of plain images (instead of just two) for which a particular linear combination gives a delta function. It is then sufficient to compute the corresponding linear combination of the ciphered images to obtain the key.

Note that Attack 5 is not applicable in the case of the phase-encoded variant of the encryption scheme [12]. This is because it is impossible to obtain zero-valued pixels in this case, thus forbidding the formation of a delta function. However, Attack 6 remains possible. For instance, let us assume that amplitude images with values between 0 and 1 are encoded to phases between 0 and 2π. Now we choose amplitude Image 1 to be uniformly zero and amplitude Image 2 to also be uniformly zero except in the center (0,0) where it takes a value of 0.5. Then the corresponding phase encoded plain images will be P_1 that is uniformly 1 (phase 0) and P_2 that will be uniformly 1 except in the center where it takes a value of -1 (phase π). The difference $P_1 - P_2$ is therefore 0 at every pixel except in the center where it take the value $1-(-1)=2$. Therefore we can see that $P_1 - P_2 = 2\delta$ and Attack 6 is applicable. It has to be noted that, since the plain images are phase coded, the knowledge of the first key X is necessary to complete the decryption process. Fortunately, this key can be recovered with a single additional plain-ciphered image pair.

Attack 7: Employ Attack 5 or 6, and further obtain the ciphered image corresponding to a uniform image as plain image; deduce the first key X.

Indeed, if P_3 is a uniform (spatially constant) image, with its corresponding ciphered image C_3, then Eq. (2) gives

$$P_3 \cdot X = \mathscr{F}^{-1}\left(C_3 \div Y\right), \tag{4}$$

and since P_3 is constant, the first key is directly recovered as

$$X \propto \mathscr{F}^{-1}\left(C_3 \div Y\right). \tag{5}$$

To sum up, with at most three chosen plain-ciphered image pairs, it is possible to recover the two encryption keys and break the system.

5. Known plaintext attacks

We now assume that the attacker knows several plain images with their corresponding ciphered images. The plain images do not have to have very specific forms so they do not have to be chosen by the attacker.

5.1. Vectorial notation

For the following analysis, we consider that P, C, X and Y are discrete (pixelized) and spatially bounded. This is always the case in practice, given that even continuous images and phase keys can be adequately sampled. We therefore represent these four two-dimensional functions by four arrays. The total number of pixels of each array is N. We re-arrange each array into a vector with N components; this vector is formed by sequentially appending each column of the original array. The four vectors are called P, C, X and Y. The linearity of the encryption

#80383 - $15.00 USD Received 23 Feb 2007; revised 27 Apr 2007; accepted 10 May 2007; published 30 Jul 2007

scheme (see Section 4) means that the transformation between the plain image vector P and the ciphered image vector C can be represented by a matrix-vector product as follows

$$\mathsf{C} = \mathbf{T}\mathsf{P}, \tag{6}$$

where $\mathbf{T}$ is the $N \times N$ transformation matrix that characterizes the encryption process.

5.2. *Using a base of the vector space*

As a linear operation in a vector space of dimension N, the encryption is fully defined by its operation on a base of the vector space (that is, a set of N linearly independent vectors).

Attack 8: Obtain the ciphered images corresponding to N known linearly independent plain images; retrieve the key by matrix inversion.

An $N \times N$ matrix $\mathbf{P}$ can be formed, using as columns the vectors $\mathsf{P}_1, \ldots \mathsf{P}_N$ that correspond to the N plain images. Similarly, an $N \times N$ matrix $\mathbf{C}$ can be formed with the vectors corresponding to the ciphered images. The process of encrypting all the plain images can be written as

$$\mathbf{C} = \mathbf{T}\mathbf{P}, \tag{7}$$

where $\mathbf{T}$ is the same encryption matrix as in Eq. (6). Now, $\mathbf{P}$ and $\mathbf{C}$ are known to the attacker, and $\mathbf{P}$ is invertible because its columns are linearly independent. So, the encryption matrix can be computed by

$$\mathbf{T} = \mathbf{C}\mathbf{P}^{-1}, \tag{8}$$

with $\mathbf{P}^{-1}$ the inverse of $\mathbf{P}$. At this stage the system is compromised because knowing the encryption matrix $\mathbf{T}$ is equivalent to knowing the keys. Indeed, from Eq. (6), we see that any unknown plain image P can be obtained from the ciphered image C by

$$\mathsf{P} = \mathbf{T}^{-1}\mathsf{C}, \tag{9}$$

where $\mathbf{T}^{-1}$ is the inverse of $\mathbf{T}$.

Note that the attack presented in this subsection is quite general and would work for any linear encryption technique [30, 31]. In practice, this attack is not very useful because it requires one to know a huge number of plain images, for instance 10000 images if each image has 100×100 pixels.

5.3. *Using two known images*

Here we assume that the attacker knows two plain images P and P' with their corresponding ciphered images C and C'.

Attack 9: Obtain the ciphered images corresponding to 2 known plain images and deduce the keys by solving a linear system of equations.

Let us express the encryption operation in terms of p_i, p'_i, c_i, c'_i, x_i and y_i, the elements of the vectors P, P', C, C', X and Y, respectively. For the first image, we can write

$$c_i = y_i \sum_{j=1}^{N} F_{ij} x_j p_j \quad \text{with} \quad 1 \leq i \leq N, \tag{10}$$

#80383 - $15.00 USD Received 23 Feb 2007; revised 27 Apr 2007; accepted 10 May 2007; published 30 Jul 2007

where F_{ij} is the matrix that characterizes the FT operation. Note that, although it operates on vectors, it does not represent a one-dimensional FT but rather a two-dimensional FT on the images before they are re-arranged in vector form. For the second image, we can write similarly

$$c'_i = y_i \sum_{j=1}^{N} F_{ij} x_j p'_j \quad \text{with} \quad 1 \leq i \leq N. \tag{11}$$

Now, it is possible to eliminate the elements of the second key y_i by cross-multiplying Eqs. (10) and (11). We get

$$c_i \sum_{j=1}^{N} F_{ij} x_j p'_j = c'_i \sum_{j=1}^{N} F_{ij} x_j p_j \quad \text{with} \quad 1 \leq i \leq N, \tag{12}$$

hence

$$\sum_{j=1}^{N} F_{ij} x_j (c_i p'_j - c'_i p_j) = 0 \quad \text{with} \quad 1 \leq i \leq N. \tag{13}$$

In Eq. (13), the only unknown variables are the components of the first key x_j. We thus rewrite this equation as

$$\sum_{j=1}^{N} S_{ij} x_j = 0 \quad \text{with} \quad 1 \leq i \leq N, \tag{14}$$

where $S_{ij} = F_{ij}(c_i p'_j - c'_i p_j)$ is known. Equation (14) represents a system of N linear equations with N unknown variables. This system has a trivial solution $x_i = 0$ for every i. This solution is not the one we are looking for because we know that the first key is not zero. Therefore there has to be another solution to this system. This is only possible if the determinant of the matrix $\mathbf{S}$ is zero, that is if the N equations are not linearly independent. The interpretation of this property is that the keys are defined up to a constant phase factor. We can thus arbitrarily choose the value of one pixel of the key and this will fix the values of the rest of the pixels. We decide to fix $x_N = 1$. We can also eliminate the last equation of the system, because it is a linear combination of the other equations. With these modifications, Eq. (14) is transformed to

$$\sum_{j=1}^{N-1} S_{ij} x_j = -S_{iN} \quad \text{with} \quad 1 \leq i \leq N-1. \tag{15}$$

This new system of $N-1$ equations with $N-1$ variables can be solved by classical system solving techniques such as Gauss elimination or LU decomposition [32]. Of course, the resolution is only possible if the determinant of this new system is not zero. In practice, this condition is not very restrictive and it is easy to find two images for which it holds. Once the first key is known, the second key is easily retrieved using for instance Eq. (10) by

$$y_i = \frac{c_i}{\sum_{j=1}^{N} F_{ij} x_j p_j} \quad \text{with} \quad 1 \leq i \leq N. \tag{16}$$

An example of this attack is presented in Fig. 5. The known plain images P and P' are shown in Figs. 5(a) and 5(b) respectively. From these two images and their corresponding ciphered images, the keys X and Y are found. These recovered keys are then used to decrypt the unknown ciphered image shown in Fig. 5(c). The result of the decryption is given in Fig. 5(d). As can be seen, the plain image is successfully decrypted.

#80383 - $15.00 USD Received 23 Feb 2007; revised 27 Apr 2007; accepted 10 May 2007; published 30 Jul 2007

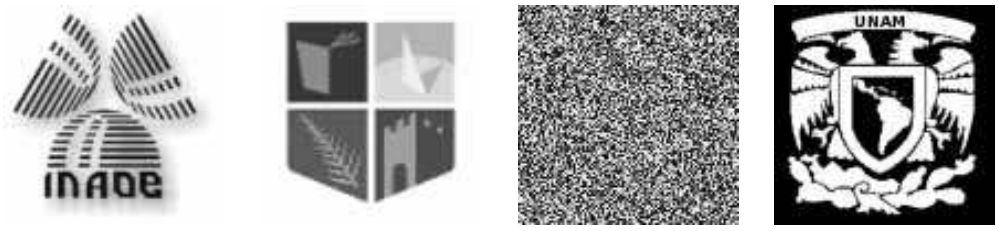

Fig. 5. Decryption using Attack 9. (a) and (b) are the two known plain images. (c) Ciphered image corresponding to an unknown plain image. (d) Unknown image decrypted using the keys retrieved by Attack 9.

The attack presented in this section requires solving a system of about N equations. The complexity of this attack is thus $O(N^2)$ in space and $O(N^3)$ in time. This polynomial complexity makes the problem tractable. For the 100×100 images of Fig. 5, the keys were found in 2 hours 15 minutes on a desktop computer (Pentium 4 HT, 3.0 GHz, 2 GB RAM) using Gaussian elimination with backsubstitution.

Note that all the variables involved in the attack can be complex-valued. The only restrictions are that all x_i, all y_i and a sufficient number of p_i and p'_i be different from 0. This situation means in particular that p_i can take complex values as happens in the phase-encoded encryption. The attack therefore works without any modification in this latter case. It would even be applicable to a mixed amplitude-phase encoding. Also, the encryption keys need not be pure phase to be recovered.

5.4. Effect of noise

In all the attacks described above, we have assumed that the ciphered images are known without error. This might not be the case in practice. In this section, we assume that the ciphered images are known up to an additive gaussian noise. We have already said that Attacks 1-4 are not practical. For Attack 5-7, it is not difficult to see that the error on the retrieved key is proportional to the noise on the ciphered images. Attacks 8 and 9 are more heavily affected by noise since the linear system resolutions involved in the key recovery process amplify the errors. Let us focus on the most interesting case of Attack 9. Figure 6(a) presents six plain images, of which the two leftmost are supposed to be known. The corresponding ciphered images are degraded by a 5% additive gaussian noise. We apply Eqs. (15) and (16) to recover the keys from the two known images, as in Section 5.3. Figure 6(b) shows the images decrypted with the recovered keys. It can be seen that, even with this low noise, the images are not recognizable, except maybe the ones that were used to compute the keys.

In order to more accurately retrieve the encryption keys in spite of the noise, we need to use more than two plain-ciphered image pairs. We now assume that we know four such pairs, although the process can be generalized to any number of pairs. The four known plain images P^1, P^2, P^3, P^4 are the four leftmost images in Fig. 6(a). For any two pairs of known plain-ciphered images (P^α, C^α) and (P^β, C^β), we can obtain a system of equations identical to Eq. (14):

$$\sum_{j=1}^{N} S_{ij}^{\alpha\beta} x_j = 0 \quad \text{with} \quad 1 \leq i \leq N, \tag{17}$$

where $S_{ij}^{\alpha\beta} = F_{ij}(c_i^\alpha p_j^\beta - c_i^\beta p_j^\alpha)$ and p_j^α, p_j^β, c_i^α and c_i^β are the pixels of P^α, P^β, C^α and C^β

#80383 - $15.00 USD Received 23 Feb 2007; revised 27 Apr 2007; accepted 10 May 2007; published 30 Jul 2007

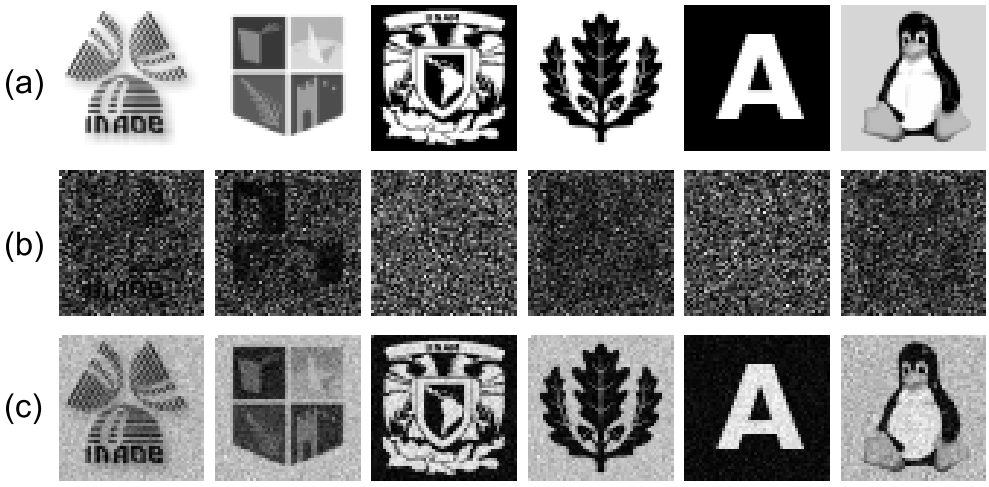

Fig. 6. Decryption of noisy ciphered images using an adaptation of Attack 9. (a) Plain images. (b) Decrypted images using the keys retrieved by simple system solving. (c) Decrypted images using the keys retrieved by least-square solving.

respectively. Equation (17) can also be written in matrix form as

$$\mathbf{S}^{\alpha\beta}\mathsf{X} = 0, \tag{18}$$

where $\mathbf{S}^{\alpha\beta}$ is the $N \times N$ matrix that contains the $S_{ij}^{\alpha\beta}$, X is a $N \times 1$ vector that contains the unknown pixels of the first key, and the right-hand term is a $N \times 1$ vector of zeros. Now we can associate the four known images in six different ways in order to obtain six different versions of Eq. (18) with varying values of α and β. These six equations, together with the condition $x_N = 1$ that avoids the trivial solution $\mathsf{X} = 0$ (see Section 5.3), can be combined into a single equation:

$$\begin{bmatrix} & \mathbf{S}^{12} & \\ & \mathbf{S}^{13} & \\ & \mathbf{S}^{14} & \\ & \mathbf{S}^{23} & \\ & \mathbf{S}^{24} & \\ & \mathbf{S}^{34} & \\ 0 & \dots \quad 0 & 1 \end{bmatrix} \mathsf{X} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}, \tag{19}$$

where the left-hand term is the product of a $(6N+1) \times N$ matrix by a $N \times 1$ vector, and the right-hand term is a $(6N+1) \times 1$ vector. Equation (19) defines an overdetermined system of equations. A least-square solution of this system can be found by classical techniques such a QR decomposition or singular value decomposition [32]. The obtained key X is still noisy; in particular it is not pure phase as it is supposed to be. We can enforce this property by normalizing the amplitude of each pixel:

$$\hat{x}_i = \frac{x_i}{|x_i|} \quad \text{with} \quad 1 \le i \le N. \tag{20}$$

We now need to recover the second key Y from the first (normalized) key $\hat{X}$. Rather than a direct derivation from a single image using Eq. (16), it is more accurate to find again a least-

#80383 - $15.00 USD Received 23 Feb 2007; revised 27 Apr 2007; accepted 10 May 2007; published 30 Jul 2007

square solution using the four known images. Since the $\hat{x}_i$ are now known, Eq. (10) for image α can be rewritten as

$$\lambda_i^\alpha y_i = c_i^\alpha \quad \text{with} \quad 1 \leq i \leq N, \tag{21}$$

where

$$\lambda_i^\alpha = \sum_{j=1}^{N} F_{ij} \hat{x}_j p_j^\alpha. \tag{22}$$

Equation (21) can be written in the following matrix form

$$\Lambda^\alpha \mathsf{Y} = \mathsf{C}^\alpha, \tag{23}$$

where Λ^α is a diagonal matrix given by

$$\Lambda^\alpha = \begin{bmatrix} \lambda_1^\alpha & & \\ & \ddots & \\ & & \lambda_N^\alpha \end{bmatrix}. \tag{24}$$

The four known images provide four different versions of Eq. (23) that can be combined into a single equation:

$$\begin{bmatrix} \Lambda^1 \\ \Lambda^2 \\ \Lambda^3 \\ \Lambda^4 \end{bmatrix} \mathsf{Y} = \begin{bmatrix} \mathsf{C}^1 \\ \mathsf{C}^2 \\ \mathsf{C}^3 \\ \mathsf{C}^4 \end{bmatrix}. \tag{25}$$

Again, Eq. (25) can be solved in a least-square sense by classical algorithms and the obtained Y can be normalized by

$$\hat{y}_i = \frac{y_i}{|y_i|} \quad \text{with} \quad 1 \leq i \leq N. \tag{26}$$

Figure 6(c) presents the images decrypted using the retrieved keys $\hat{X}$ and $\hat{Y}$. The results are now very satisfactory, including for the two rightmost images that were not used in the key recovery process. These results demonstrate that a good approximation of the keys was found. If the noise were stronger, it would be necessary to incorporate more known images in the key estimation process.

6. Discussion

This paper presents an analysis of the theoretical — rather than practical — security of the DRP technique. The various presented attacks assume an abstract version of the technique where all data is available in digital form. In a real optical implementation, the difficulty to manipulate and measure the involved physical quantities adds some challenges to the attacker.

We have shown that Attacks 1 through 4 are intractable, which makes the DRP encryption resistant to brute force attacks. Attack 8 requires a very large number of known plain images and is therefore quite impractical. On the other hand, Attacks 5 to 7 are surprisingly easy and effective if the attacker is able to choose images to be encrypted. If this possibility does not exist, Attack 9 is still very powerful as a known plain image attack which only requires two known images. The DRP technique must therefore be considered potentially insecure against chosen and known plaintext attacks. Its main weakness lies in its linearity.

#80383 - $15.00 USD Received 23 Feb 2007; revised 27 Apr 2007; accepted 10 May 2007; published 30 Jul 2007

The attacks presented in this paper have some advantages over other attacks proposed recently. For instance, the attack described in [29] requires access to an encryption machine and uses a large number of chosen plain images, which makes it impractical. Similarly, the technique in [25] requires access to a decryption machine and a large number of chosen ciphered images. In addition, the latter paper mentions an attack with a centered delta function as input image. In this paper we have shown that a non-centered delta function will also suffice, as will any set of innocuous images whose linear combination is a shifted delta function. This latter attack defeats the defense proposed by Carnicer et al. [25]. Further, we have shown that with the possibility of a second chosen image (an image with constant complex amplitude) the first phase mask can also be found, allowing complex-valued inputs to be decrypted.

The attacks presented in [27] and [28] use heuristic approaches and result in an approximated recovery of the keys and the decrypted images are quite noisy. Actually, Ref. [28] does not show the decryption of an unknown image. On the other hand, our Attacks 5 to 9 are based on a mathematical analysis of the DRP technique and they are able to recover the exact keys and therefore provide a perfect decryption of unknown images. Additionally, it is important to note that the attacks proposed in [25, 27, 28, 29] only work for the case of amplitude-coded plain images. In our proposal, Attacks 6 through 9 are perfectly applicable to phase-encoded images so they can be used against the full phase DRP encryption [12].

A variant of the DRP encoding that uses keys in the Fresnel domain has been proposed in [14]. Optical encoding using keys in the Fresnel domain will be substantially more secure as any attack requires a search in 3D domain to locate the precise Fresnel location of the keys. Therefore, various attacks presented in the literature which use the linearity property will not be very practical. Also, other encryption techniques based on fractional Fourier transform have been proposed [8, 16, 18, 23]. The most effective attacks presented in this paper are not directly applicable to these variant schemes using Fresnel domain keys and fractional Fourier transform. These variant schemes should therefore be preferred to the original DRP encryption.

7. Conclusion

In this paper we have described several ways to attack the double random phase encryption technique, some of which are impractical and other are very effective. We have shown that an exhaustive search of the key is generally intractable, even when applying some simplifications to reduce the number of combinations. However, we have presented chosen and known plaintext attacks that are able to efficiently recover the keys of the system. The most dangerous attack only requires two known plain images. We have studied the effects of noise on this attack and proposed a way to reduce them. We have also provided a comparison of the proposed attacks to other techniques described in the literature. Given the risks involved with the presented attacks, it is recommendable to be extremely cautious when using the double random phase encryption. Large keys of at least 1000×1000 pixels should be used and, if possible, it should be avoided to re-use the same keys for different images, as in a one-time pad approach. Also, a safer alternative to the original double random phase encryption is to use variant schemes such as keys in the Fresnel domain or fractional Fourier transform encryption.

Acknowledgments

The authors are grateful to the anonymous reviewers for their useful comments. Y. Frauel, A. Castro and T. J. Naughton acknowledge support from Enterprise Ireland and Science Foundation Ireland.

#80383 - $15.00 USD Received 23 Feb 2007; revised 27 Apr 2007; accepted 10 May 2007; published 30 Jul 2007