Optics Communications 281 (2008) 5122–5125

Contents lists available at ScienceDirect

Optics Communications

journal homepage: www.elsevier.com/locate/optcom

Collision in double random phase encoding

Guohai Situ [a,*], David S. Monaghan [b], Thomas J. Naughton [c,d], John T. Sheridan [b], Giancarlo Pedrini [a], Wolfgang Osten [a]

[a] *Institut für Technische Optik, Universität Stuttgart, 70569 Stuttgart, Germany*

[b] *School of Electrical, Electronic and Mechanical Engineering, College of Engineering, Mathematics and Physical Sciences, Optoelectronic Research Centre, The SFI-Strategic Research Cluster in Solar Energy Conversion, University College Dublin, Belfield, Dublin 4, Ireland*

[c] *Department of Computer Science, National University of Ireland, Maynooth, County Kildare, Ireland*

[d] *University of Oulu, RFMedia Laboratory, Oulu Southern Institute, Vierimaantie 5, 84100 Ylivieska, Finland*

ARTICLE INFO

Article history:
Received 23 April 2008
Received in revised form 2 July 2008
Accepted 9 July 2008

PACS:
42.30.Kq
42.30.Va
42.30.Rx

Keywords:
Double random phase encoding
Collision
Cryptanalysis
Phase retrieval

ABSTRACT

Collision is a situation that occurs when two or more distinct inputs into a security system produce identical outputs, which is undesirable in some security applications. This is especially true of the applications of watermarking and authentication. In this manuscript we present a study of the collision property of double random phase encoding. We show that one can produce meaningful collisions from the cyphertext of a watermark embedded in a host image by use of phase retrieval techniques.

1. Introduction

The term *collision* refers to an event in which two or more bodies strike or come together in physics. However, in terms of information security, it is a situation that occurs when two or more distinct inputs into a security system produce identical outputs. Collision is undesirable in some security applications. For example in the case of the hash function, which produces a unique fixed-size string (hash code) from any input and is widely used in data integrity and password verifications techniques. If an attacker finds an input which will produce a hash code identical to a legal one, he can access the system as a legal user. Similarly, any watermarking system should also have the property of collision resistance. Otherwise, it is impossible to distinguish the ownership of any watermarked property.

Double random phase encoding (DRPE) has been used for optical encryption, image hiding and watermarking [1–3]. Recently there have been several studies pointing out that random phase encoding systems are vulnerable to chosen-cyphertext, chosen-plaintext and known-plaintext attacks [4–8]. However, to date there is no study on the collision resistance of this technique. So our motivation in this manuscript is to study the collision property of DRPE.

2. Collision algorithm

In DRPE, an input image or watermark, $f(x,y)$, is encoded into a stationary white noise, i.e., the cyphertext $g(x,y)$, using two statistically independent random phase distributions $\phi(x,y)$ and $\psi(u,v)$ located in a 4f system at the input and Fourier planes, respectively. For image hiding and watermarking applications, $g(x,y)$ is then embedded into a host image to yield a watermarked image [2,3]. The host image does not have much effect to our analysis because of the reason stated in the next paragraph. Here we just focus on phase encoding in our discussions. The encoding process can be mathematically expressed as [1]

$$g(x,y) = \mathscr{F}\{\mathscr{F}\{f(x,y)\exp[\mathrm{j}\phi(x,y)]\}\exp[\mathrm{j}\psi(u,v)]\}, \quad (1)$$

where (x,y) and (u,v) are the coordinates of the spatial and Fourier domains, respectively, and the symbol $\mathscr{F}\{\cdot\}$ represents Fourier transform. The two random phase distributions $\phi(x,y)$ and $\psi(u,v)$,

[*] Corresponding author.
E-mail address: ghsitu@gmail.com (G. Situ).

0030-4018/$ - see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.optcom.2008.07.011

or only the latter in the case of real input, act as security keys to the system. In general it is difficult to recover the watermark, $f(x,y)$, directly from the cyphertext $g(x,y)$ without using these two keys. The purpose of the cryptanalysis presented in [4–8] is to retrieve part, or all, of the information of these keys with *a priori* but incomplete knowledge of the plaintext and/or the cyphertext, and eventually retrieve any plaintext encrypted using the same key-set $[\phi(x,y), \psi(u,v)]$. Collision, on the other hand, involves finding another key-set, $[\phi'(x,y), \psi'(u,v)]$, which will encrypt a different $f'(x,y)$ to the same $g(x,y)$, the cyphertext of $f(x,y)$. If such a collision exists, the illegal user can claim the ownership of any property, such as digital art, watermarked with $g(x,y)$, which is legally owned by another person.

To address this problem, we assume that the illegal user can access the full complex cyphertext $g(x,y)$. This assumption is reasonable because in some watermarking systems based on random phase encoding, $g(x,y)$ is directly added into a host image to obtain the watermarked image [2]. If the host image is available as in some applications [9], it is easy to extract the cyphtertext. In the cases when the host image and $g(x,y)$ are not separable, it does not affect the performance of the following algorithm since the spectrum of the watermarked image can then be used as the constraint in the Fourier plane. So we can neglect the host image in our discussions. With $g(x,y)$ one could also obtain its full complex Fourier transform:

$$G(u,v) = |G(u,v)|\exp[j\beta(u,v)] = \mathscr{F}^{-1}\{g(x,y)\}. \tag{2}$$

In this case the collision problem can be stated as: Find a real image, $f'(x,y)$, and two phase distributions, $\exp[j\phi'(x,y)]$ and $\exp[j\psi'(u,v)]$, so that the following equation is valid:

$$|G(u,v)|\exp[j\beta(u,v)] = \mathscr{F}\{f'(x,y)\exp[j\phi'(x,y)]\}\exp[j\psi'(u,v)]. \tag{3}$$

An illegal user who intends to act as a legal user has complete flexibility in choosing a self-defined real image as the input collision $f'(x,y)$. Once such an input image is chosen, the amplitudes on both sides of Eq. (3) are determined. The aim then reduces to finding the phase $\exp[j\phi'(x,y)]$, or equivalently $\exp[j\psi'(u,v)]$, given the two intensity measurements, respectively, in the spatial and Fourier planes. This problem can be solved using an iterative Fourier transform (IFT) algorithm [10]. Assuming the algorithm has reached the $k^{\rm th}$ iteration and retrieved the phase $\exp[j\phi'_k(x,y)]$, one can form an estimate of the image in the spatial domain $f'_k(x,y) = |f'(x,y)|\exp[j\phi'_k(x,y)]$. The succeeding iteration then consists of the following steps:

(I) Fourier transform the estimate of the image:

$$F_k(u,v) = |F_k(u,v)|\exp[j\alpha_k(u,v)] = \mathscr{F}\{f'_k(x,y)\}; \tag{4}$$

(II) Replace the modulus of the resulting spectrum, $|F_k(u,v)|$, by $|G(u,v)|$ to form the estimate of the Fourier transform:

$$F'_k(u,v) = |G(u,v)|\exp[j\alpha_k(u,v)]; \tag{5}$$

(III) Inverse Fourier transform the modulated complex amplitude back to the spatial domain:

$$\hat{f}'_{k+1}(x,y) = |\hat{f}'_{k+1}(x,y)|\exp[j\phi'_{k+1}(x,y)] = \mathscr{F}^{-1}\{F'_k(u,v)\}; \tag{6}$$

(IV) Replace the resulting modulus, $|\hat{f}'_{k+1}(x,y)|$, by $f'(x,y)$:

$$f'_{k+1}(x,y) = f'(x,y)\exp[j\phi'_{k+1}(x,y)]. \tag{7}$$

In general there is no analytic solution for Eq. (3), but feasible solutions always exist [10]. The above iterative process is repeated until the convergent criteria are satisfied. Normally the converging factor can be chosen as the mean square error (MSE) or the correlation coefficient between $f'(x,y)$ and $|\hat{f}'(x,y)|$. Since the IFT algorithm is essentially an error-reduction algorithm, it is therefore reasonable to determine the convergence by tracking the MSE and test whether it is smaller than a predefined threshold value or not.

When the iterative process converges after N iterations, the phase distribution $\phi'(x,y) = \phi'_N(x,y)$ numerically obtained is apparently one of the keys expected. According to the description in Step II, the Fourier spectrum $|F_N(u,v)|$ of $f'(x,y)\exp[j\phi'_N(x,y)]$ should satisfy the Fourier domain constraint, that is, $|F_N(u,v)| \approx |G(u,v)|$. However the phase component $\exp[j\alpha_N(u,v)]$ may not, therefore one must add an additional phase to this to satisfy the phase constraint. This phase difference can be regarded as the key $\psi'(u,v) = \beta(u,v) - \alpha_N(u,v)$.

3. Computer simulation

We perform computer simulations to demonstrate the above theoretical analysis. The image we used as the legal watermark $f(x,y)$ is the image "elaine" of 512 × 512 pixels in size, which can be found in the USC-SIPI database [11]. Following Eq. (1) it is easy to calculate the corresponding cyphertext $g(x,y)$ using two random phase masks $\exp[j\phi(x,y)]$ and $\exp[j\psi(u,v)]$, which are white noise distributions and not shown in the text. Once the collision $f'(x,y)$ is chosen, for example the image "lena" [11], the above algorithm can be applied. The converging criteria used in the simulation is the MSE with the threshold value γ_t equal to, for example, 0.0015.

The algorithm converges within 100 iterations. The converging behavior in the logarithm scale is shown by the curve with triangle marks in Fig. 1. One can see that the error curve drops very fast in the first few iterations, and then becomes much flatter but keeps decreasing. This is the typical converging behavior of the error-reduction algorithm [10]. Two random phase distributions $\exp[j\phi'(x,y)]$ and $\exp[j\psi'(u,v)]$ were then retrieved. Generally, these two distributions are different from the legal phase keys. The phase differences between the legal and illegal phase keys, $\Delta\phi(x,y) = \mathrm{mod}\{[\phi(x,y) - \phi'(x,y)], 2\pi\}$ and $\Delta\psi(u,v) = \mathrm{mod}\{[\psi(u,v) - \psi'(u,v)], 2\pi\}$, also have random noise feature.

In our simulation, the expected values of $\Delta\phi(x,y)$ and $\Delta\psi(u,v)$ are 3.1368 and 3.1369 radians, which are close to π; and the standard deviations of both are found to be 1.8144. In order to more clearly show the random characters of these two phase maps, we plot their histograms in Fig. 2, which indicates they are uniformed distributions between $[0, 2\pi]$. Even though the retrieved illegal phase keys are significantly different from the legal ones, they

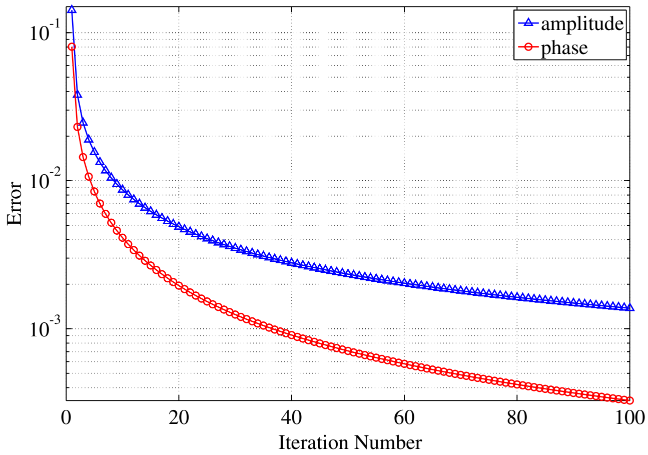

Fig. 1. The converging behavior of the algorithm in the logarithm scale. The MSE is calculated using the formula $M^{-2}\sum_{j=1}^{M}\sum_{i=1}^{M}[I(i,j) - |I'_k(i,j)|]^2$, where $I(i,j)$ and $I'_k(i,j)$ are the spatial constraint and the retrieved image at the kth iteration, and $M \times M$ is the size of the image. As explained in the text, this plot shows the behavior of the algorithm for both amplitude-encoded and phase-encoded input schemes.

can be used to decode the chosen image $f'(x,y)$ from $g(x,y)$ with high quality as shown in Fig. 3, the MSE between which and “lena” is 0.0015, and the normalized correlation coefficient between them is 0.9911.

Note that we did not impose any restriction on the selection of the real collision image $f'(x,y)$, the initialization of the algorithm, or the spectrum of the cyphertext in the above discussions and simulations. The convergence property holds for all the applications of the error-reduction algorithms as quoted from [10]. What this means is that the illegal user has a lot of flexibility in choosing $f'(x,y)$; and he can always retrieve two feasible keys $\exp[\mathrm{j}\phi'(x,y)]$ and $\exp[\mathrm{j}\psi'(u,v)]$ from the chosen image and the obtained Fourier spectrum. Therefore the algorithm converges regardless of whose Fourier spectrum, $g(x,y)$, in the case of encryption, or the weighted summation of $g(x,y)+$ a host image, in the case of watermarking, being used. This property also results in very serious problems in practical applications: anybody who can access $g(x,y)$ can claim his or her ownership of anything watermarked by $g(x,y)$. Once many such claim the ownership of an identical image involving specific $f'(x,y)$ and phase key-pair, it would become difficult for a moderator to conduct a technical assessment regarding property rights.

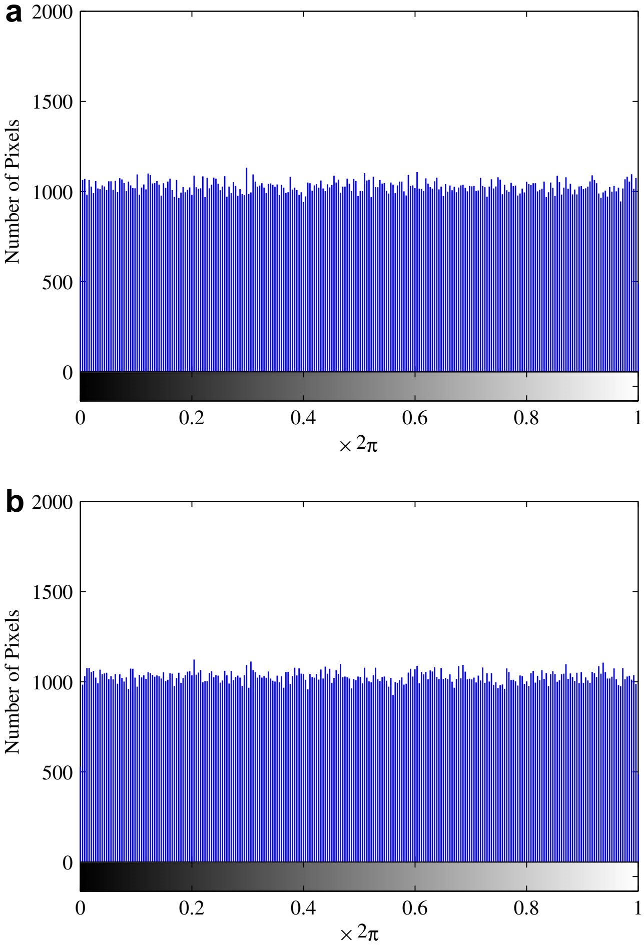

Fig. 2. The histograms of the phase differences between the legal and illegal phase keys: (a) $\mathrm{mod}\{[\phi(x,y)-\phi'(x,y)],2\pi\}$ and (b) $\mathrm{mod}\{[\psi(u,v)-\psi'(u,v)],2\pi\}$.

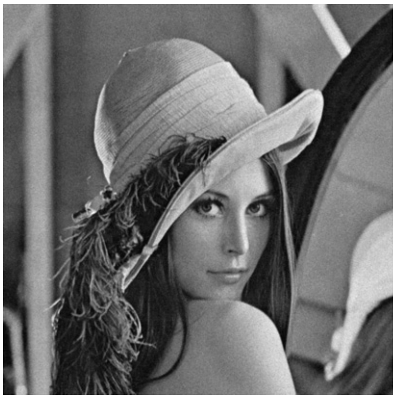

Fig. 3. The high-quality collision image retrieved in the simulations.

Even if the watermark $f(x,y)$ is pre-encoded into a phase-only function, $\exp[\mathrm{j}2\pi f(x,y)]$, it is still possible to find collisions although such pre-processing is helpful in enhancing the difficulty of cryptanalysis, as is also the case for encryption [12]. In this case, one just needs to change the constraint in the spatial domain (Step IV in the algorithm) to be a unity amplitude. In this study two random-like phase distributions $\exp[\mathrm{j}\phi'(x,y)]$ and $\exp[\mathrm{j}\psi'(u,v)]$ can always be found so that the algorithm converges [10]. It is shown in Fig. 1 that the algorithm converges faster due to the unity amplitude of the spatial constraints in this case. Then $\phi'(x,y)$ can be written as the summation of any collision image $2\pi f'(x,y)$ and a random-like distribution, i.e., $\exp[\mathrm{j}\phi'(x,y)]=\exp\{\mathrm{j}[2\pi f'(x,y)+\phi''(x,y)]\}$. It is $\exp[\mathrm{j}\phi''(x,y)]$ and $\exp[\mathrm{j}\psi'(u,v)]$ that now serve as the illegal keys.

4. Conclusion

Cryptographic hash functions are used for authentication (digital signatures such as watermarks), message integrity, and password verification. Collisions in such functions pose serious security risks. While there is no collision risk with DRPE when the key is not part of the input (by definition, because it is a symmetric encryption technique), we have shown that DRPE admits collisions when the key is regarded as an input variable and DRPE is used as a cryptographic hash function. Although it is a necessary condition for ideal encryption [14], this property represents a weakness that needs to be fixed for DRPE to be used for authentication in this way.

The existence of collisions arises because of the linear nature of the system and the methodology of phase encoding. Although the introduction of additional detection system in the Fourier domain may be helpful in allowing a moderator to assess the ownership [4,13] when it is necessary, some information about the spectrum embedded in the watermarked image is then necessary. This additional information however may be utilized by the illegal user to analyze the watermark and the keys.

Acknowledgements

G. Situ gratefully acknowledges the support of the Alexander von Humboldt Foundation. T. Naughton acknowledges support from a Marie Curie Fellowship from the European Commission Framework Programme 6. This work is also supported by Science Foundation Ireland, Enterprise Ireland and the Irish Research Council for Science, Engineering and Technology.

References

[1] P. Réfrégier, B. Javidi, Opt. Lett. 20 (1995) 767.
[2] S. Kishk, B. Javidi, Appl. Opt. 11 (2002) 5462.
[3] S. Kishk, B. Javidi, Opt. Express 11 (2003) 874.
[4] A. Carnicer, M. Montes-Usategui, S. Arcos, I. Juvells, Opt. Lett. 30 (2005) 1644.
[5] U. Gopinathan, D.S. Monaghan, T.J. Naughton, J.T. Sheridan, Opt. Express 14 (2006) 3181.
[6] X. Peng, P. Zhang, H. Wei, B. Yu, Opt. Lett. 31 (2006) 1044.

[7] G. Situ, U. Gopinathan, D.S. Monaghan, J.T. Sheridan, Appl. Opt. 46 (2007) 5257.
[8] Y. Frauel, A. Castro, T.J. Naughton, B. Javidi, Opt. Express 15 (2007) 10253.
[9] I.J. Cox, M.L. Miller, J.A. Bloom, Digital Watermarking, Morgan Kaufmann, 2002.
[10] J.R. Fienup, Appl. Opt. 21 (1982) 2758.
[11] <http://sipi.usc.edu/database/>.
[12] N. Towghi, B. Javidi, Z. Luo, J. Opt. Soc. Am. A 16 (1999) 1915.
[13] G. Situ, J. Zhang, Opt. Commun. 245 (2005) 55.
[14] C.E. Shannon, Bell Syst. Tech. J. 28 (1949) 656.