Image Encryption System Based on a Nonlinear Joint Transform Correlator for the Simultaneous Authentication of Two Users

Juan M. Vilardy O. 1,*, María S. Millán 2 and Elisabet Pérez-Cabré 2

1 Grupo de Óptica e Informática, Department of Electronic Engineering, Universidad Popular del Cesar, Valledupar (Cesar) 200001, Colombia
2 Applied Optics and Image Processing Group, Universitat Politècnica de Catalunya · BarcelonaTech, 08222 Terrassa (Barcelona), Spain; m.millan@upc.edu (M.S.M.); elisabet.perez@upc.edu (E.P.-C.)

* Correspondence: vilardy.juan@unicesar.edu.co

Received: 31 October 2019; Accepted: 12 December 2019; Published: 14 December 2019

Abstract: We propose a new encryption system based on a nonlinear joint transform correlator (JTC) using the information of two biometrics (one digital fingerprint for each user) as security keys of the encryption system. In order to perform the decryption and authentication in a proper way, it is necessary to have the two digital fingerprints from the respective users whose simultaneous authentication is pursued. The proposed security system is developed in the Fourier domain. The nonlinearity of the JTC along with the five security keys given by the three random phase masks and the two digital fingerprints of the two users allow an increase of the system security against brute force and plaintext attacks. The feasibility and validity of this proposal is demonstrated using digital fingerprints as biometrics in numerical experiments.

Keywords: image encryption; joint transform correlator (JTC); authentication

1. Introduction

The field of image encryption using optical and optoelectronic devices has shown a great potential to protect the information contained in these images [1–4]. One of the most important techniques for the image encryption using optical devices is the technique of double random phase encoding (DRPE) proposed by Réfrégier y Javidi [5]. The DRPE uses two random phase masks (RPMs) placed in the input and the Fourier plane, respectively, of the encryption system in order to convert an original image in a stationary white noise image (encrypted image). The DRPE can be optically implemented using a 4f-processor [6] or a joint transform correlator (JTC) [7]. The initial JTC architecture was modified in the works of [8–11] with the purpose of simplifying the optical setup of the encryption system, but this modification affected the decrypted images with a poor image quality [12]. We introduced a nonlinear modification of the JTC architecture that allowed an enhancement of the decrypted image quality and an improvement over the security of the encrypted image [12]. The JTC used in [7–12] was performed in the Fourier domain (FD) and it was extended from the FD to the Fresnel domain [13–17], the fractional Fourier domain [18–21] and the Gyrator domain [22]. The nonlinear modifications of the JTC architecture introduced in [13,14,18,19,22] allowed the retrieval of the correct decrypted images and an improved resistance of the encryption system against brute force and plaintext attacks.

Some DRPE-based systems have been proposed for multifactor authentication [23–25]. These optically inspired systems combine DRPE encryption, matched filter, and photon-counting techniques for simultaneous pattern recognition of up to four signals. The distortion-invariant ID tags were used in [24,26] to obtain encryption and authentication systems with scale and rotation invariance. The procedure described for the distortion-invariant ID tags in [24,26] could be applied to the proposed
encryption and authentication system to overcome, at least to some extent, the possible problems that would arise if the information to inspect was affected by scale and rotation changes.

In this work, we present a DRPE-based encryption-decryption system that also incorporates an authentication stage. The proposed system uses a nonlinear JTC and five security keys consisting of three RPMs and two images to authenticate. When all five keys are used correctly in the decryption process, the decrypted code reveals the successful authentication of the two images. Otherwise, the validation fails. This proposal can be implemented when access to restricted areas must be controlled with a higher level of security. In some occasions, two users may access coordinately to the same place, or if there is only one person, two different biometrics can be considered as further restricted level of security. Another situation where this simultaneous authentication can be applied will be the case of identifying simultaneously a person and a vehicle or package. Random phase codes are generated in the encryption stage, for instance being designed to allow the entrance of a particular place or in a particular date. This idea was already mentioned in [23–25], but its implementation followed a different procedure. The encryption algorithm was digital and none of the encrypted signals were retrieved but just authenticated.

The novelty of this proposal compared with our previous works [12–14,18,19,22], can be described in three different aspects. Firstly, the security system incorporates an authentication stage of the encrypted signals. Secondly, the proposed system has in total five keys consisting of three RPMs and two digital fingerprints of the users (two biometric images) increasing the system security. Finally, the applied nonlinearity of the proposed system differs from previous proposals in order to achieved a satisfactory decryption and authentication of the whole set of signals.

The triple random-phase encoding (TRPE) is the DRPE implemented with a 4f-processor that uses one additional RPM at the output plane of the encryption system [27]. The proposed encryption system in this work also uses three RPMs, but the DRPE technique is performed using a nonlinear JTC architecture. In addition, the three RPMs for the TRPE are placed at the input, Fourier, and output planes of the encryption system whereas the three RPMs for the proposed encryption system of this work are placed at the input plane of the JTC. A new attack based on deep learning shows that the DRPE and TRPE implemented with a 4f-processor are vulnerable [28]. The DRPE and TRPE implemented using a 4f-processor are linear systems and the attack proposed in [28] is limited to images (plaintexts or images to encrypt) encoded in amplitude (real-valued). This attack based on deep learning should be adapted to nonlinear systems and images to encrypt encoded in phase (complex-valued) in order to succeed when attacking the proposed security system of this work [12,13,18,22].

2. Encryption, Decryption, and Authentication Stages

2.1. Encryption Stage

The image to encrypt \( f(x, y) \) has only two real values: 0 or 1 (binary image). The image \( f(x, y) \) is a code whose decryption reveals the positive validation of the two images introduced for authentication.

The following images are RPMs

\[

to f(x, y) = \exp\{i2\pi s(x, y)\}, \quad k_1(x, y) = \exp\{i2\pi n_1(x, y)\}, \quad k_2(x, y) = \exp\{i2\pi n_2(x, y)\},
\]

where \( x \) and \( y \) are the spatial coordinates, and \( s(x, y) \), \( n_1(x, y) \) and \( n_2(x, y) \) are normalized positive function generated randomly, statistically independent, and uniformly distributed with values in the interval of \([0, 1]\) [5]. All the functions used in the encryption and decryption stages are images with \( M \times N \) pixel size. Figure 1a shows the optical encryption scheme (part I) using a nonlinear JTC architecture and the optical decryption scheme (part II) based on two successive Fourier transform (4f-processors).
Figure 1. (a) The scheme of the optical setup composed of an encryption system based on a nonlinear JTC architecture (part I) and a decryption system based on a 4f-processor (part II). Distributions contained in the input plane of the (b) encryption system, and (c) decryption system.

The input plane of the JTC is composed by three data distributions spatially separated as it is depicted in Figure 1b. The first data distribution is the function \( g(x,y) \), which is determined by the original image to encrypt \( f(x,y) \) encoded in phase \( f_{ph}(x,y) = \exp\{i2\pi f(x,y)\} \) [29] multiplied by the RPM \( r(x,y) \)

\[
g(x,y) = f_{ph}(x,y)r(x,y) = \exp\{i2\pi[f(x,y) + s(x,y)]\}. \tag{2}
\]

The second data distribution of the input plane of the JTC is the function \( h_1(x,y) \), which is the product of RPM \( k_1(x,y) \) and the image of the digital fingerprint of the first user \( p_1(x,y) \) encoded in phase

\[
h_1(x,y) = \exp\{i2\pi p_1(x,y)\}k_1(x,y) = \exp\{i2\pi[p_1(x,y) + n_1(x,y)]\}. \tag{3}
\]

The third data distribution of the input plane of the JTC is the function \( h_2(x,y) \), given by the product of RPM \( k_2(x,y) \) by the image of the digital fingerprint of the second user \( p_2(x,y) \) encoded in phase

\[
h_2(x,y) = \exp\{i2\pi p_2(x,y)\}k_2(x,y) = \exp\{i2\pi[p_2(x,y) + n_2(x,y)]\}. \tag{4}
\]
The three data distributions of the input plane are phase-only distributions, which are located as it is indicated in Figure 1b. The distribution $g(x, y)$ is placed centred at the coordinates $(x, y) = (a, b)$. The second and third data distributions are centred at the coordinates $(x, y) = (-a, b)$ and $(x, y) = (-a, -b)$, respectively. We assume that all three data distributions do not overlap spatially; that is, they are nonzero in a spatial region that is not larger than $2a \times 2b$ (and preferably smaller to avoid border effects). The intensity of the Fourier transform of the input plane of the JTC, which is the joint power spectrum (JPS), is given by

$$
\text{JPS}(u, v) = |\mathcal{F}\{g(x-a, y-b) + h_1(x+a, y-b) + h_2(x+a, y+b)\}|^2
$$

where $x$ and $y$ are the spatial frequency coordinates, the functions represented by capital letters correspond to the Fourier transforms of the functions represented in lowercase letters.

The following intensities $|G(u, v)|^2$, $|H_1(u, v)|^2$, and $|H_2(u, v)|^2$, which can be sequentially captured by displaying different input planes, are subtracted from the JPS. Then, the previous result is divided by the nonlinear term $|H_1(u, v)|^2 + |H_2(u, v)|^2$ in order to obtain the encrypted image $E(u, v)$

$$
E(u, v) = \frac{\text{JPS}(u, v) - \left(|G(u, v)|^2 + |H_1(u, v)|^2 + |H_2(u, v)|^2\right)}{|H_1(u, v)|^2 + |H_2(u, v)|^2}
$$

$$
= \frac{1}{T(u, v)} \left[ G(u, v)H_1^*(u, v)e^{-i2\pi(2au)} + G(u, v)H_2^*(u, v)e^{-i2\pi(2bu)} + G^*(u, v)H_1(u, v)e^{-i2\pi(-2au)} + H_1(u, v)H_2^*(u, v)e^{-i2\pi(2bu)} + G^*(u, v)H_2(u, v)e^{-i2\pi(-2bu)} + H_2(u, v)H_2(u, v)e^{-i2\pi(-2bu)}\right],
$$

where $T(u, v) = |H_1(u, v)|^2 + |H_2(u, v)|^2$ and the superscript $*$ denotes the complex conjugation operation. The encrypted image $E(u, v)$ has real values and it is obtained from four intensities distributions. The five security keys of the encryption system are given by the three RPMs ($r(x, y)$, $k_1(x, y)$, and $k_2(x, y)$) and the two images of the digital fingerprints of the two users $p_1(x, y)$ and $p_2(x, y)$.

All the steps concerning Fourier transformations described in the encryption and decryption procedures can be performed optically by considering the optical processor shown in Figure 1a. Optical processing has the valuable property of inherent parallelism, which allows for fast encryption of large volumes of data. In addition to this, the security strength of optical cryptography resides in the ability of optics to process the information in a hyperspace of states, where variables such as amplitude, phase, polarization, wavelength, spatial position, and fractional spatial frequency domain can all be used to hide the signal with greater concealment [1]. Image subtraction and nonlinearity are applied to the camera acquired intensity distributions by digital computation.

### 2.2. Decryption and Authentication Stages

In the decryption stage, the users willing to be authenticated provide their personal fingerprints in-situ. These fingerprints are compared with the encrypted data, which for instance, can be attached on an ID card.

The processor obtains the necessary random phase codes from a database, which can vary depending on the area to access, the date, or other particular information. Once, all the required information is available for the processor, the following step of the decryption system (Figure 1a, part II) is to multiply the encrypted image $E(u, v)$ by the Fourier transform of the input plane of the decryption system (Figure 1c), and the result is
\[
D(u, v) = E(u, v) \mathcal{F} \{ h_1(x + a, y - b) + h_2(x + a, y + b) \} \\
= E(u, v) \left[ H_1(u, v)e^{-i2\pi(-au+bv)} + H_2(u, v)e^{-i2\pi(-au-bv)} \right] \\
= \frac{1}{T(u, v)} \left[ G(u, v)|H_1(u, v)|^2e^{-i2\pi(au+bv)} + G(u, v)H_1(u, v)H_2^*(u, v)e^{-i2\pi(au+3bv)} \right. \\
+ G^*(u, v)H_1^2(u, v)e^{-i2\pi(-3au+bv)} + H_1^2(u, v)H_2^*(u, v)e^{-i2\pi(-au+3bv)} + G(u, v)H_1^2(u, v)H_2(u, v)e^{-i2\pi(-au-bv)} \right. \\
+ G(u, v)H_1(u, v)H_2(u, v)e^{-i2\pi(-3au-bv)} + H_1(u, v)H_2^*(u, v)e^{-i2\pi(-au+3bv)} + G^*(u, v)H_1(u, v)H_2^*(u, v)e^{-i2\pi(-au+3bv)} \\
+ G^*(u, v)H_2^2(u, v)e^{-i2\pi(-3au-bv)} + H_2^2(u, v)H_2^*(u, v)e^{-i2\pi(-au-3bv)} \right]. \tag{7}
\]

The inverse Fourier transform of Equation (7) provides the output plane of the 4f-processor. Different distributions are spatially separated at the output plane. For decryption purposes, only the first and eighth terms of Equation (7) are the most relevant information in order to recover the original image that was encrypted. The distributions in the output plane corresponding to these two terms cover the same spatial region, centred at coordinates \((a, b)\); the other terms, however, contribute to the output plane with spatially separated distributions from them. Thus, considering only the first and eighth term of Equation (7) to compute the relevant distributions of the output plane, the spatial region centred at \((a, b)\) is

\[
d_{18}(x, y) = \mathcal{F}^{-1} \left\{ \frac{G(u, v) \left( |H_1(u, v)|^2 + |H_2(u, v)|^2 \right) e^{-i2\pi(au+bv)}}{|H_1(u, v)|^2 + |H_2(u, v)|^2} \right\} = g(x - a, y - b). \tag{8}
\]

The decrypted image is obtained from the function \(d_{18}(x, y)\) as follows

\[
2\pi f(x - a, y - b) = \arg \{ d_{18}(x, y) r^*(x - a, y - b) \}, \tag{9}
\]

where \(\arg\) is the phase of a complex-valued function. If the five security keys used in the decryption system are the same security keys employed in the encryption system, the decrypted image \(f(x, y)\) is a replica of the original image \(f(x, y)\) utilized in the encryption system. The correct retrieval of the original image in the decryption process is achieved through the nonlinear modifications introduced in the IFs when the encrypted image \(E(u, v)\) is computed.

Finally, the disclosure of the successful authentication is done over the decrypted image \(f(x, y)\). If any of the five security keys used in the decryption system is different from the five security keys in the encryption system, the decrypted image \(f(x, y)\) will not be a replica of the original image \(f(x, y)\) and therefore, the two users will not be authenticated.

3. Computational Simulations

The computational simulations of the security system proposed in Section 2 are presented in Figure 2. The images utilized in the security system have a resolution of 300 \times 300 pixels and, in general, are grayscale. The original image \(f(x, y)\) and the random code image \(s(x, y)\) of the RPM \(r(x, y)\) are shown in Figure 2a,b, respectively. The random code images \(n_1(x, y)\) and \(n_2(x, y)\) of the RPMs \(k_1(x, y)\) and \(k_2(x, y)\), respectively, have different values but the same appearance of the image presented in Figure 2b. The images of the digital fingerprints of the two users are presented in Figure 2c,d, which correspond to the functions \(p_1(x, y)\) and \(p_2(x, y)\), respectively.

The encrypted image \(E(u, v)\) is shown in Figure 2e. This encrypted image has a noisy appearance that neither reveals any information of the original image \(f(x, y)\) nor the fingerprints of the users. If the decryption process is performed using the encrypted image \(E(u, v)\) and the five security keys \((r(x, y), k_1(x, y), k_2(x, y), p_1(x, y), p_2(x, y))\) with their correct values, the original image \(f(x, y)\)
will be recovered ideally at the output of the decryption system. The decrypted image \( \hat{f}(x, y) \) obtained from the encrypted image \( E(u, v) \) and the correct values of the five security keys, is shown in Figure 2f.

![QR Code](image1.png)

**Figure 2.** (a) Original image to encrypt \( f(x, y) \). (b) Random code image \( s(x, y) \) of the random phase mask (RPM) \( r(x, y) \). Images of the digital fingerprints of the two users: (c) first user \( p_1(x, y) \) and (d) second user \( p_2(x, y) \). (e) Encrypted image \( E(u, v) \). (f) Decrypted image \( \hat{f}(x, y) \) using the correct five security keys \( r(x, y), k_1(x, y), k_2(x, y), p_1(x, y) \) and \( p_2(x, y) \). Decrypted images for the following wrong security keys: (g) the image of the digital fingerprint of the first user \( p_1(x, y) \) and (h) the RPM \( k_2(x, y) \).

In this work, to evaluate the quality of the decrypted images, we use the metric of the root mean square error (RMSE) between the decrypted image \( \hat{f}(x, y) \) and the original image \( f(x, y) \) [12]

\[
\text{RMSE} = \sqrt{\frac{\sum_{x=1}^{M} \sum_{y=1}^{N} [f(x, y) - \hat{f}(x, y)]^2}{\sum_{x=1}^{M} \sum_{y=1}^{N} [f(x, y)]^2}},
\]

where \( M = N = 300 \). The RMSE metric that evaluates the quality of the decrypted image takes values in \([0,1]\), when the value of the RMSE metric is near or equal to 0, this metric indicates an excellent quality of the image for the retrieval of the decrypted image at the output of the decryption system, whereas the values of the RMSE metric near or equal to 1 represent a worse quality of the decrypted image. The RMSE between the original image of Figure 2a and the decrypted image of Figure 2f is 0.057.

The decrypted image is presented in Figure 2g when a wrong image of the digital fingerprint of the first user \( p_1(x, y) \) is used in the decryption system. If the RPM \( k_2(x, y) \) is wrong in the decryption system, the decrypted image obtained is shown in Figure 2h. The values of the RMSEs between the original image of Figure 2a and the decrypted images of Figure 2g,h are 0.87 and 0.89, respectively.

Finally, the decrypted image \( \hat{f}(x, y) \) of Figure 2f (which is a replica of the original image presented in Figure 2a), is read as a QR code obtaining the successful authentication and therefore, the simultaneous authentication of the two users using the two images of the digital fingerprints in
the decryption process is performed in a correct way. For the images of Figure 2g,h, the operation of authentication fails because at least one of the five security keys is wrong for the decryption process.

The key space analysis for the proposed security system consists of every possible combination of the security keys: the three RPMs \((r(x, y), k_1(x, y), \text{ and } k_2(x, y))\) and the two images of the digital fingerprints of the two users \(p_1(x, y)\) and \(p_2(x, y)\). Each security key has a resolution of \(300 \times 300\) pixels in grayscale and each pixel has 256 possible values. A rough estimation of the number of attempts required to retrieve the three RPMs and the two images of the digital fingerprints of the two users, is of the order of \(256^{5(300)\times(300)} = 256^{450,000}\). Therefore, the brute force attacks are intractable just considering all the possibilities of the five keys of the proposed security system [30]. The nonlinear modifications applied on the JPS allow a better protection against the chosen-plaintext and known-plaintext attacks than the linear cryptosystems, as it was demonstrated in references [12,13,18,22].

4. Conclusions

We have presented a DRPE based encryption-authentication system using a nonlinear JTC architecture, where the biometric information of two users are employed as security keys. The proposed security system has five security keys represented by three RPMs and two images of digital fingerprints of two users. These five security keys along with the nonlinearities introduced over the JTC are intended to improve the security of the encrypted image against brute force and plaintext attacks. It is necessary to have the five security keys previously mentioned, in order to obtain the right decrypted image and to perform the authentication process in a correct way. The authentication process can be performed, whenever the resulting image (decrypted image) of the decryption process is a replica of the original image that was initially encrypted. The excellent quality for the decrypted image is due to the nonlinear modifications introduced on the JTC architecture. Finally, the encryption and decryption systems based on a JTC allow the simultaneous authentication of two users using the biometric information of these two users.

Author Contributions: The work described in this article was the collaborative development of all authors. Conceptualization, J.M.V.O., M.S.M. and E.P.-C.; methodology, J.M.V.O., M.S.M. and E.P.-C.; software, J.M.V.O.; validation, M.S.M. and E.P.-C.; investigation, J.M.V.O., M.S.M. and E.P.-C.; writing—original draft preparation, J.M.V.O.; writing—review and editing, M.S.M. and E.P.-C.; supervision, M.S.M. and E.P.-C.

Funding: This research has been funded by the Universidad Popular del Cesar from Valledupar (Cesar), Colombia, and the Spanish Ministerio de Ciencia e Innovación and Fondos FEDER (Project DPI2016-76019-R).

Conflicts of Interest: The authors declare no conflicts of interest.

References


