Information Sciences 222 (2013) 323–342

Contents lists available at SciVerse ScienceDirect

Information Sciences

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

Local Shannon entropy measure with statistical tests for image randomness

Yue Wu [a,*], Yicong Zhou [b], George Saveriades [a], Sos Agaian [c], Joseph P. Noonan [a], Premkumar Natarajan [d]

[a] Department of Electrical and Computer Engineering, Tufts University, 161 College Ave., Medford, MA 02155, USA
[b] Department of Computer and Information Science, University of Macau, Ave. Padre Tomás Pereira, Taipa, Macau, China
[c] Department of Electrical and Computer Engineering, University of Texas at San Antonio, One UTSA Circle, San Antonio, TX 78249, USA
[d] Raytheon BBN Technologies, 10 Moulton Street, Cambridge, MA 02138, USA

ARTICLE INFO

Article history:
Received 24 March 2011
Received in revised form 25 July 2012
Accepted 30 July 2012
Available online 7 August 2012

Keywords:
Image encryption
Shannon entropy
Image randomness
Hypothesis test

ABSTRACT

In this paper we propose a new image randomness measure using Shannon entropy over local image blocks. The proposed local Shannon entropy measure overcomes several weaknesses of the conventional global Shannon entropy measure, including unfair randomness comparisons between images of different sizes, failure to discern image randomness before and after image shuffling, and possible inaccurate scores for synthesized images. Statistical tests pertinent to this new measure are also derived. This new measure is therefore both quantitative and qualitative. The parameters in the local Shannon entropy measure are further optimized for a better capture of local image randomness. The estimated statistics and observed distribution from 50,000 experiments match the theoretical ones. Finally, two examples are given, applying the proposed measure to image randomness among shuffled images and encrypted images. Both examples show that the proposed method is more effective and more accurate than the global Shannon entropy measure.

1. Introduction

Since the information age began in the late 1970s, the digital world has kept evolving on a nearly daily basis. More particularly, the past 10 years have seen an impressive growth of capabilities of electronic devices as well as their usage in virtually all walks of life (i.e. smartphones, digital music players, home robotic devices, electronic readers, etc.). These devices highlight the fast increase in computational and storage facilities of modern electronics. Compared to the rapid development of electronic devices and computer computational capacities, contemporary data encryption technologies are not very different from those of 10 years ago: many data encryption algorithms in use 10 years ago are still in use today, such as the data encryption standard (DES) [2] dating from 1976, the Blowfish cipher [7] from 1993, the Twofish [46] cipher from 1998, and the advanced encryption standard (AES) [3] from 1998. Although shortcomings of these methods on bulk data, such as digital images and digital videos, have been pointed out [61], these old algorithms still dominate encryption methods at all levels (individuals, organizations, companies and governments).

Image encryption has recently become a fertile research area. Many new image encryption algorithms or methods have been proposed, e.g. chaotic system based image ciphers [6,12,15,16,22,28,29,37,40,42,54,55,63,64], SCAN language based algorithms [13,14], transform based algorithms [35,38,49,60,65]. The goal of image encryption is to turn an input image, commonly referred to as *plaintext*, into an unrecognized or unintelligent image, referred to as *ciphertext*, using a predetermined method, which is called an *image cipher*. For adversaries without plaintext knowledge, an image cipher works like

[*] Corresponding author. Tel.: +1 617 627 3217.
E-mail address: ywu03@ece.tufts.edu (Y. Wu).

0020-0255/$ - see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.ins.2012.07.049

a symbol source generating pixels in ciphertext. Since the completely random source achieving the maximum randomness has a uniform distribution, it is desirable that an image cipher has an indistinguishable distribution. Otherwise, the image cipher is insecure as patterns can be identified through the observation of a sufficiently large number of encrypted images [11,32,33,56].

Image randomness can be measured using a variety of methods, such as histogram analysis [6,12,16,22,28,29,35,37,40,42,54,55,57,60,63,64], global Shannon entropy measure [6,12,42,55,64,66], adjacent pixel correlations [6,12,16,35,42,57,64,66]. One major drawback of these conventional techniques is that they provide quantitative rather than qualitative measures. However, it is qualitative measures that make it is possible to distinguish patterned data from random-like data. In contrast, many statistical tests that provide quantitative measures (e.g. the Kolmogorov test [50], poker test [50], gap test [50], autocorrelation test [50], diffusion randomness test [27], and available test suites such as FIPS 140-1 [1] and 140-2 [4]) are designed for either a stream or a block cipher, which is built for one dimensional bit-stream rather than a two dimensional image. These methods are therefore not directly applicable to image data.

Although the importance of statistical tests for image randomness is obvious, little work has been done on this particular topic. Refs. [34,62] discussed a randomness measure defined on image edges; however, it is still a quantitative measure which only gives a randomness score. Statistical tests for the number of changing pixel rate (NPCR) and the unified average changing intensity (UACI) [6,12,16,35,54,55,57,64,66], two measurements on the changing rate of encrypted images, have been proposed recently in [59], but they are made for testing randomness between two images rather than on the randomness of one image.

In this paper, we develop new statistical tests for image randomness based on the local Shannon entropy measure, which is a generalization of conventional Shannon entropy. The remainder of the paper is organized as follows: Section 2 gives a brief review on Shannon entropy, the central limit theorem and random number generators in cryptography; Section 3 introduces a random image generator model, and derives the mean and variance of Shannon entropy for random images; Section 4 defines the local Shannon entropy measure and statistical tests for random images, and optimizes parameters of the local Shannon entropy measure to attain the best localization capacity; Section 5 compares the theoretical statistics and distributions with those observed from a large scale simulation with 50,000 random images; Section 6 presents possible applications of the proposed method for image shuffling and image encryption; and Section 7 concludes the paper.

2. Preliminaries

2.1. Shannon entropy measure and properties

Shannon entropy [47], named after Claude Shannon, was first proposed in 1948. Since then, Shannon entropy has been widely used in the information sciences. Shannon entropy is a measure of the uncertainty associated with a random variable. Specifically, Shannon entropy quantifies the expected value of the information contained in a message. The Shannon entropy of a random variable X can be defined as in Eq. (1), where P_i is defined in Eq. (2) with x_i indicating the ith possible value of X out of n symbols, and P_i denoting the possibility of $X = x_i$.

$$\mathrm{H}(X) = \mathrm{H}(P_1, \dots, P_n) = -\sum_{i=1}^{n} P_i \log_2 P_i \tag{1}$$

$$P_i = \Pr(X = x_i) \tag{2}$$

Shannon Entropy attains, but is not limited to, the following properties:

(a) Bounded: $0 \leqslant \mathrm{H}\,(X) \leqslant \log_2 n$
(b) Symmetry: $\mathrm{H}\,(P_1, P_2, \dots) = \mathrm{H}\,(P_2, P_1, \dots) = \cdots$
(c) Grouping [45]: $\mathrm{H}\,(P_1, \dots, P_n) = \mathrm{H}\,(P_1 + P_2, P_3, \dots, P_n) + (P_1 + P_2)\,\mathrm{H}(P_1/(P_1 + P_2), P_2/(P_1 + P_2))$

In the context of digital images, an $M \times N$ image X can be interpreted as a sample from an L-intensity-scale source that emitted it. As a result, we can model the source symbol probabilities using the histogram of the image X (the observed image) and generate an estimate of the source entropy [23]. For example, an 8-bit gray image allows $L = 256$ gray scales from 0 to 255. Additionally, denote the number of pixels within image X at pixel intensity scale l as n_l. Then

$$P_l = \Pr(X = l) = n_l/T \tag{3}$$

where $l \in \{0, 1, \dots, L-1\}$ denotes the intensity scale and $T = M \times N$ is the total number of pixels in image X. Therefore, the Shannon entropy score of image X can be calculated as shown in Eq. (4).

$$\mathrm{H}(X) = -\sum_{l=1}^{L-1} P_l \log_2 P_l = \sum_{l=0}^{L-1} \frac{n_l}{T} \log_2 \frac{T}{n_l} \tag{4}$$

The theoretical maximum of the Shannon entropy score for an L symbol source is $\log_2 L$, when each symbol is equally likely distributed, i.e.

$$P_0 = P_1 = \cdots = P_l = \cdots = P_{L-2} = P_{L-1} = 1/L \tag{5}$$

Shannon entropy [8,47] has been widely used in image encryption for years as a common measure for information and uncertainty [17,21,36,39].

2.2. Central limit theorem

Let $Y_1, Y_2, \ldots, Y_n$ be a sequence of n independent and identically distributed observations on a random variable Y associated with a finite mean μ and a variance σ^2. The central limit theorem (CLT) states that the sample mean of these observations will be approximately normally distributed, with a mean μ and a variance σ^2/n, when the number of samples n is sufficiently large. Mathematically, the CLT can be stated as follows:

$$\overline{Y_n} = \sum_{i=1}^{n} \frac{Y_i}{n} \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right), \quad as\ n \to \infty \tag{6}$$

The most important merit of the CLT is that the probability density function (PDF) of $\overline{Y_n}$ is dependent on the mean and variance of the random variable Y. In other words, it is not necessary to know the exact PDF of Y for computing the PDF of $\overline{Y_n}$, as long as its mean and variance are known. Heuristically, many statisticians believe that if the sample size n is larger than 30 then it is sufficiently large [41,52], while others suggest larger sample sizes [30], for Example 100 [9].

Instead of using $\overline{Y_n}$, the random variable Z_n defined in Eq. (7) is commonly used in hypothesis tests, where $Z_n \sim \mathcal{N}(0,1)$ as $n \to \infty$. The convergence in distribution implies that Eq. (8) is held for arbitrary $z \in \mathbb{R}$, where $\Phi(z)$ is the cumulative distribution function (CDF) of $\mathcal{N}(0,1)$. Consequently, the statistical test designed on this Z statistic is called the Z-test, where $F_{Z_n}(z)$ denotes the actual CDF of Z_n

$$Z_n = \frac{\overline{Y_n} - \mu}{\sigma/\sqrt{n}} \tag{7}$$

$$\lim_{n\to\infty} \Pr(Z_n \leqslant z) = \lim_{n\to\infty} F_{Z_n}(z) = \Phi(z) \tag{8}$$

2.3. Random number generators in cryptography

A random number generator (RNG) is a computational algorithm or a physical device that generates a sequence of symbols from some known distribution, which implies that all elements in the sequence should be independent and identically distributed with a known distribution. Although the distribution of a RNG can be of any type, including Gaussian [31], Multinomial [18] or Poisson [10], the uniform distribution is prefered for RNGs in cryptography [5,20,44,51], as it makes all outcomes equally likely and implies less population information than non-uniform distributions. Many attacks or cryptanalysis methods, e.g. [11,32,33,56] can take advantage of non-uniformly distributed ciphertexts.

RNGs can be classified into two groups: pseudo RNGs (PRNGs) and true RNGs (TRNGs), where PRNGs are predetermined by a set of controllable parameters with completely predictable outputs; and TRNGs are normally dependent on certain physical phenomena, e.g. atmospheric noise [19] and Johnson noise [25], which can be considered as a set of noncontrollable parameters with completely unpredictable outputs. However, it is very difficult to distinguish a PRNG from a TRNG by observing output random number sequences [24].

A random sequence $Q = \{q_1, q_2, \ldots\}$ from a RNG (either a TRNG or a PRNG) with a uniform distribution on a finite symbol set $\mathbb{S} = \{s_1, s_2, \ldots, s_L\}$ has the following properties:

$$\Pr(q_t = s_1) = \Pr(q_t = s_2) = \cdots \Pr(q_t = s_L) = 1/L \tag{9}$$

$$\Pr(q_t|q_\tau) = \Pr(q_\tau|q_t) = 1/L \tag{10}$$

where t and τ denotes two sequence elements and $t \neq \tau$. It is easy to verify that the Shannon entropy of a RNG source with a uniform distribution on L symbols always attains the upper-bound $\log_2 L$.

Finally, it is worthwhile to note the links between a PRNG and a digital cipher (stream cipher or block cipher). A stream cipher normally combines a plaintext data stream with a key stream, which is generated from a PRNG [53]. In contrast, a block cipher itself is normally considered as a cryptographically secure PRNG [43], whose generated sequences also attain the properties Eqs. (9) and (10).

3. Shannon entropy of a random image

In this section, we first introduce a random image generator in Section 3.1. This random image generator enables us to generate random images, which share the same statistical properties as securely encrypted images, rendering them indistinguishable. In Section 3.3, we derive what would be the theoretical mean and variance of the Shannon entropy for a random image as defined in Section 3.2. Finally, this section ends with a discussion explaining the theoretically derived means and variances from Section 3.2 and the connection between those values and the preconditions of the random image generator.

3.1. A random image generator

Since a digital image can be considered as a bit/byte sequence fitting within a specific rectangle, we define a uniformly random image generator (RIG) in the following way:

Definition 1. A uniformly random image generator is a random number generator which equally likely generates an image pixel of intensity l out of L allowed intensities ($l \in \{0, 1, \ldots, L-1\}$).

Therefore, each image pixel is then generated independently with the identical uniform distribution as a random sequence element from a RNG model as defined in Eqs. (9) and (10). In other words, the following properties are held for a pixel $R(i, j)$ generated by a RIG.

$$\Pr(R(i,j) = l) = 1/L \tag{11}$$

$$\Pr(R(i,j)|R(i',j')) = 1/L \tag{12}$$

where $R(i, j)$ denotes the pixel located at the intersection of the ith row and jth column in R and $R(i', j')$ is another pixel different from $R(i, j)$. Equivalently, this indicates that a RIG is also a *memory-less* source [58]. It is noticeable that a RIG with a uniform distribution attains the maximum Shannon entropy, i.e. the maximum randomness with the Shannon entropy score $\log_2 L$.

3.2. Random images and securely encrypted images

In this paper, *a random image* is defined as an observed image generated by a RIG as defined previously. Fig. 1 shows random images of size 256 × 256 in binary, 8-bit grayscale and color formats. It can be observed that the histograms of these random images are all uniform-like, this because each image is an observation derived from a RIG.

The same relationship that exists between a PRNG and a digital cipher exists between a RIG and an image cipher. First, a secure digital cipher is supposed to attain confusion and diffusion properties [48], thus making Eqs. (11) and (12) hold for a secure image cipher. Second, an image cipher can be considered as a pseudo RIG, because it is completely controlled by cipher key(s) when the used image cipher is known. Therefore, a random image generated from a RIG should also be indistinguishable from an encrypted image generated from a secure image cipher. Fig. 2 shows securely encrypted images of size 256 × 256 in the binary, 8-bit grayscale and color formats. It can be seen that these encrypted images indeed look like random images in Fig. 1.

Consequently, a statistical test designed for testing the randomness of random images is also applicable to encrypted images, which are supposed to be statistically indistinguishable from random images if they are encrypted by a secure image cipher.

3.3. Mean and variance of Shannon entropy for a random image

Consider a random image X as a random variable. This implies the pixels in X satisfy the statistical properties of Eqs. (11) and (12).

Lemma 1. *The number of pixels at scale l out of L possible intensity scales in a random image X of size $M \times N$ follows the binomial distribution of T ($T = MN$) independent incidents with the success probability $1/L$, i.e.*

$$n_l \sim \mathcal{B}(T, 1/L) \tag{13}$$

Proof. Since any pixel x in the random image X follows the discrete uniform distribution $x \sim \mathcal{U}(0, L-1)$, i.e.

$$\Pr(x = l) = 1/L$$

we have,

$$\Pr(x \neq l) = 1 - 1/L = (L-1)/L$$

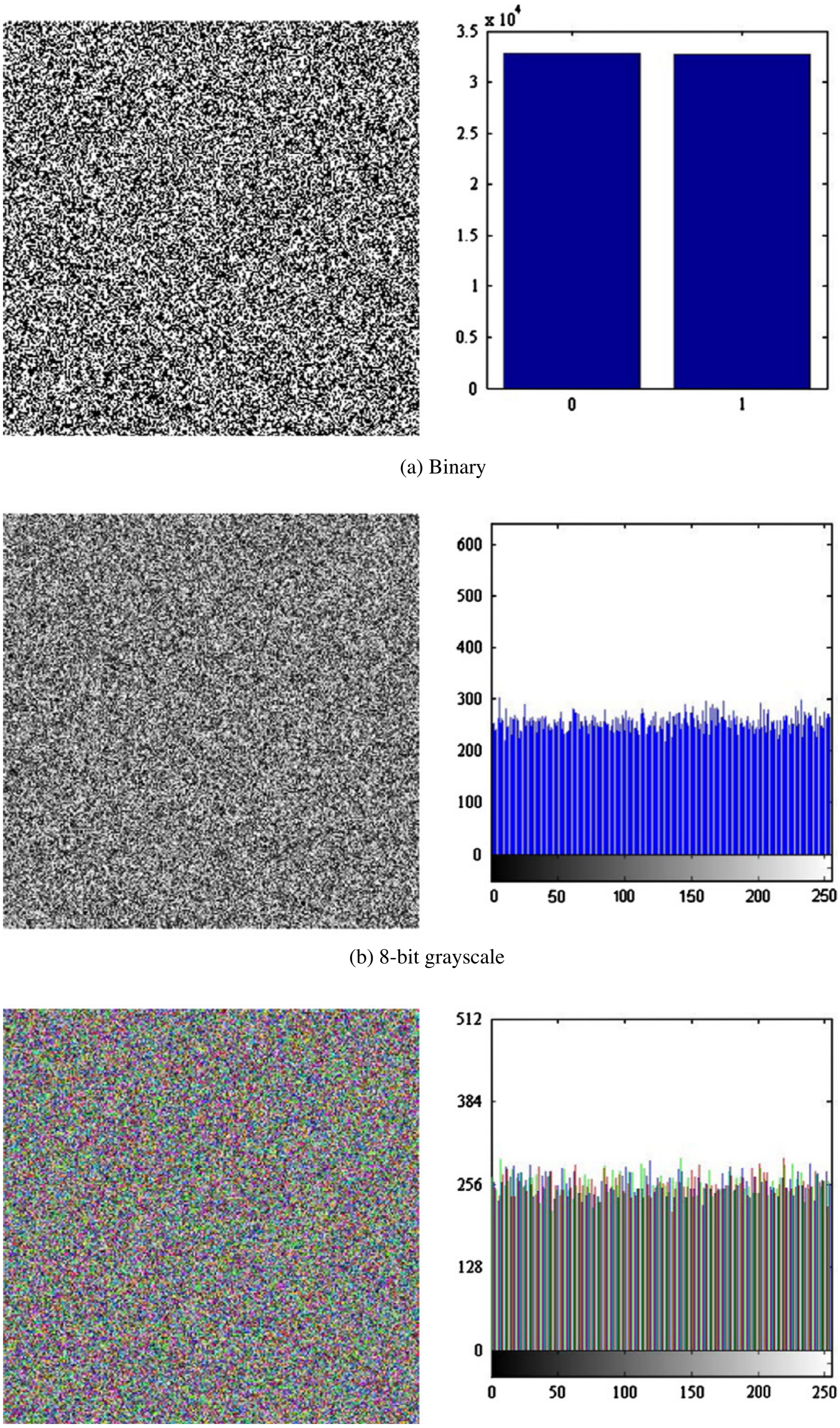
Therefore, any pixel x at intensity level l follows the Bernoulli distribution with success probability $1/L$. As a result, the number of pixels at intensity level l follows the Binomial distribution as

$$n_l \sim \mathcal{B}(T, 1/L)$$

i.e.

$$\Pr(n_l = k) = \binom{T}{k} \frac{(L-1)^{T-k}}{L^T}$$

and $\binom{T}{k} = \frac{T!}{k!(T-k)!}$ is the binomial coefficient. □

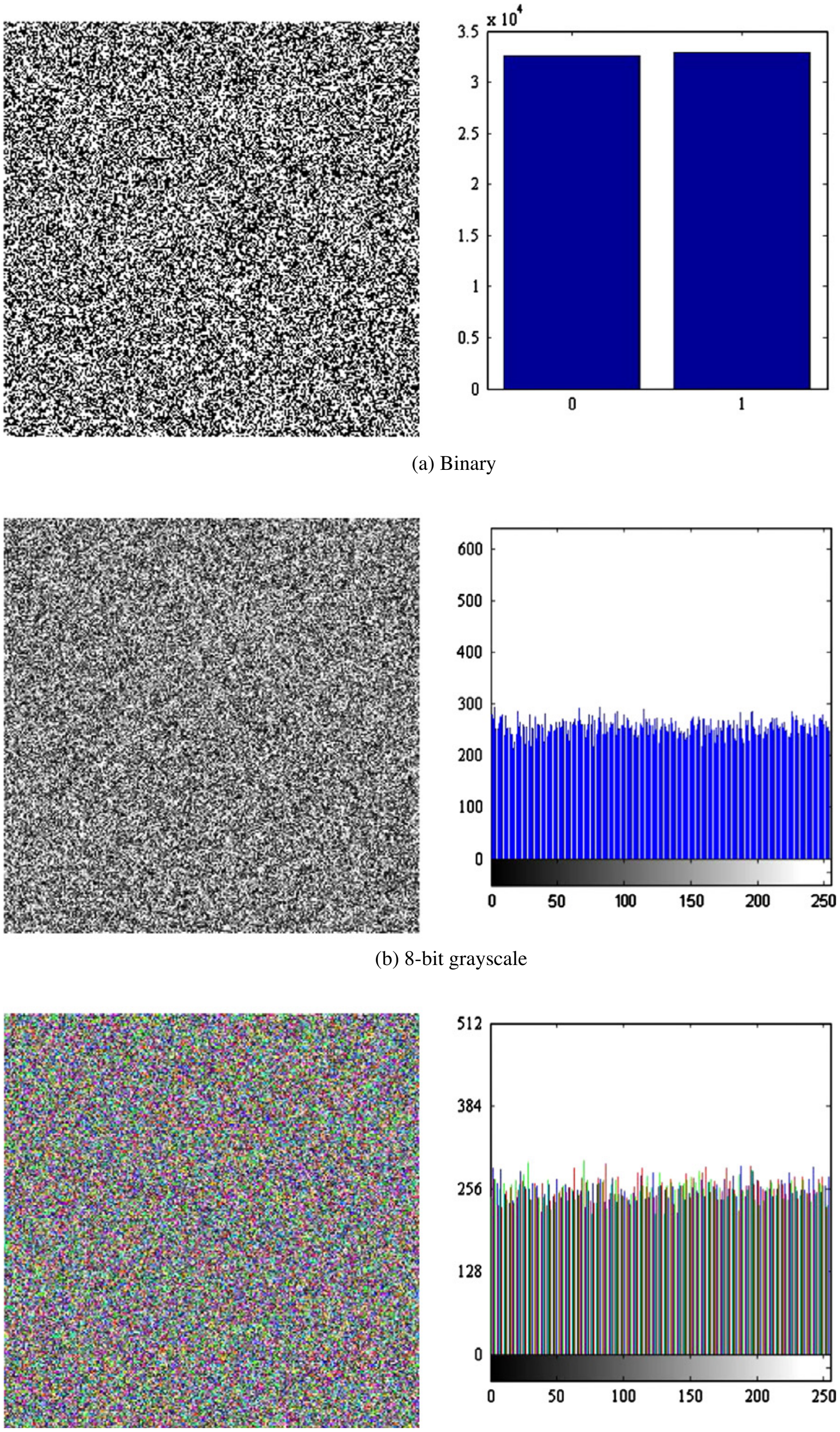


(a) Binary

(b) 8-bit grayscale

(c) Color

Fig. 1. Sample random images.



(a) Binary

(b) 8-bit grayscale

(c) Color

Fig. 2. Sample encrypted images.

Corollary 1.

$$\Pr(P_l = k/T) = \binom{T}{k}\frac{(L-1)^{T-k}}{L^T} \tag{14}$$

Proof. True since $n_l \sim \mathcal{B}(T, 1/L)$ and $P_l = n_l/T$. □

In order to find the mean and variance of the Shannon entropy of an observed random image X, i.e. $\mu_{\mathrm{H}(X)}$ and $\sigma^2_{\mathrm{H}(X)}$, we first rewrite the Shannon entropy of X as the sum of entropies from all possible intensity scales, i.e.

$$\mathrm{H}(X) = \sum_{l=0}^{L-1} h_l \tag{15}$$

where $h_l = -P_l \log_2 P_l$ denotes the Shannon entropy of intensity scale l. Using Lemma 1 that $n_l \sim \mathcal{B}(T, 1/L)$, the following statistics can be obtained:

$$\mathrm{E}[h(l)] = \mathrm{E}\Big[-\frac{n_l}{T}\log_2\frac{n_l}{T}\Big] = \sum_{n_l=0}^{T}\frac{n_l}{T}\log_2\frac{T}{n_l}\cdot\binom{T}{n_l}\frac{(L-1)^{T-n_l}}{L^T} \tag{16}$$

$$\mathrm{E}[h(l)^2] = \sum_{n_l=0}^{T}\left(\frac{n_l}{T}\log_2\frac{T}{n_l}\right)^2\cdot\binom{T}{n_l}\frac{(L-1)^{T-n_l}}{L^T} \tag{17}$$

$$\mathrm{E}[h(l_a)h(l_b)] = \sum_{n_a=0}^{T}\sum_{n_b=0}^{T-n_a}\left(\frac{n_a}{T}\log_2\frac{T}{n_a}\right)\left(\frac{n_b}{T}\log_2\frac{T}{n_b}\right)\cdot\frac{T!(L-2)^{T-n_a-n_b}}{n_a!n_b!(T-n_a-n_b)!L^T} \tag{18}$$

Subsequently, the mean $\mu_{\mathrm{H}(X)}$ and variance $\sigma^2_{\mathrm{H}(X)}$ can be derived:

$$\mu_{\mathrm{H}(X)} = \mathrm{E}[\mathrm{H}(X)] = \mathrm{E}\left[\sum_{l=0}^{L-1}h(l)\right] = \sum_{l=0}^{L-1}\mathrm{E}[h(l)] = L\cdot\mathrm{E}[h(l)] \tag{19}$$

$$\mathrm{E}[\mathrm{H}(X)^2] = \mathrm{E}\left[\left(\sum_{l=0}^{L-1}h(l)\right)^2\right] = \mathrm{E}\left[\sum_{l=0}^{L-1}h(l)^2 + \sum_{l_a=0}^{L-1}\sum_{\substack{l_b=0\\ l_b\neq l_a}}^{L-1}h(l_a)h(l_b)\right] = L\cdot\mathrm{E}[h(l)^2] + L(L-1)\cdot\mathrm{E}[h(l_a)h(l_b)] \tag{20}$$

$$\sigma^2_{\mathrm{H}(X)} = \mathrm{E}[\mathrm{H}(X)^2] - (\mathrm{E}[\mathrm{H}(X)])^2 = L\cdot\mathrm{E}[h(l)^2] + L(L-1)\cdot\mathrm{E}[h(l_a)h(l_b)] - L^2\cdot(\mathrm{E}[h(l)])^2 \tag{21}$$

Numerical results for $\mu_{\mathrm{H}(X)}$ and $\sigma_{\mathrm{H}(X)}$ using different parameter sets are shown in Table 1. It is clear from Table 1 that for a random image, $\mu_{\mathrm{H}(X)}$ increases as L increases when T is fixed; $\mu_{\mathrm{H}(X)}$ also increases as T increases when L is fixed. Meanwhile, $\sigma_{\mathrm{H}(X)}$ gets smaller as T increases.

Finally we want to point out that Eqs. (19) and (21) pave the way to the local Shannon entropy measure and tests which will be discussed in the next section, although they only reveal the mean and variance of H (X) without the exact PDF of H (X). However, as stated in the CLT, the sample mean of Shannon entropy over n random images, which is the local Shannon entropy measure, follows the Normal distribution with mean $\mu_{H(X)}$ and variance $\sigma^2_{H(X)}/n$, as long as n is sufficiently large ($n \geqslant 30$).

3.4. Discussion

It is worthwhile to understand the relationship between the derived Shannon entropy statistics of a random image and the preconditions of a random image (see Eqs. (11) and (12)).

First, the Shannon entropy of a random image is an estimation on that of the RIG which generates this random image. Because a RIG here is an L-symbol source generating equally likely symbols for each pixel, the Shannon entropy of such a RIG reaches the maximum randomness $\log_2 L$. Consequently, it is not surprising that H (X), the Shannon entropy of a random image X from this RIG, and also an estimation of the Shannon entropy for the RIG, always have a score close to the Shannon entropy of the source, $\log_2 L$.

Second, a random image with a larger image size better estimates the source entropy, and thus a random variable of H (X) has a larger mean and a smaller variance as the size of X increases. This is because a larger observed image means more pixels, and thus more symbols observed from the image source, the RIG. According to the CLT, the probability of seeing each symbol from this RIG is then better estimated in the sense that these probabilities approach their true values closer as the number of image pixels increases. Correspondingly, the estimated source Shannon entropy gets closer to the true source

Table 1
Theoretical mean and standard deviation of Shannon entropy score for a random image.

$L = 2$	Binary image		$L = 256$	Grayscale image		$L = 256$	Color image	
T	$\mu_{\mathrm{H}(X)}$	$\sigma_{\mathrm{H}(X)}$	T	$\mu_{\mathrm{H}(X)}$	$\sigma_{\mathrm{H}(X)}$	T	$\mu_{\mathrm{H}(X)}$	$\sigma_{\mathrm{H}(X)}$
2 × 2	0.780639062	0.307715375	2 × 2	1.988300234	0.076064119	2 × 2 × 3	3.542339666	0.082640020
4 × 4	0.953361607	0.066187807	4 × 4	3.942064617	0.082851351	4 × 4 × 3	5.407984610	0.079041305
8 × 8	0.988638975	0.016069236	8 × 8	5.765716929	0.076603439	8 × 8 × 3	6.938975236	0.059295884
16 × 16	0.997176704	0.003992777	16 × 16	7.174966353	0.052437999	16 × 16 × 3	7.737771412	0.023253559
32 × 32	0.999295215	0.000996718	32 × 32	7.808756571	0.017246343	32 × 32 × 3	7.939203149	0.005393141
64 × 64	0.999823868	0.000249088	64 × 64	7.954588734	0.004024888	64 × 64 × 3	7.984977322	0.001330526

Shannon entropy $\log_2 L$ and thus leads to the mean of H (X) approaching closer to $\log_2 L$ with a smaller variances. This phenomenon can be confirmed with Table 1, where $\mu_{\mathrm{H}(X)}$ increases and $\sigma_{\mathrm{H}(X)}$ decreases as T the number of pixels increases.

Finally, the derived results hold for any image pulled from an image source that satisfies the preconditions of Eqs. (11) and (12). Since an encrypted image from a secure cipher is supposed to satisfy these preconditions while an encrypted image from an insecure cipher is not, we can differentiate encrypted images derived from secure and insecure ciphers using the results derived above. More details about this application will be presented in the following sections.

4. Local Shannon entropy measure and statistical tests

Conventionally, in the image encryption community, the usage of Shannon entropy for image randomness is to compute Eq. (4) for a sample image S. This method has been widely adopted in testing the performance of an image cipher [6,12,42,55,64,66]. In this paper, this conventional usage is referred to as the *global Shannon entropy*. In contrast, the proposed Shannon entropy measure is referred to as the *local Shannon entropy*, as it only relies on a series of local blocks in an image.

4.1. Local Shannon entropy measure

We define the (k, T_B)-local Shannon entropy measure with respect to local image blocks using the following method:

- Step 1. Randomly select non-overlapping image blocks $S_1, S_2, \ldots, S_k$ with T_B pixels within a test image S of L intensity scales
- Step 2. For all $i \in \{1, 2, \ldots, k\}$ compute Shannon entropy H (S_k) via Eq. (4)
- Step 3. Calculate the sample mean of Shannon entropy over these k image blocks $S_1, S_2, \ldots, S_k$ via Eq. (22)

$$\overline{H_{k,T_B}}(S) = \sum_{i=1}^{k} \frac{\mathrm{H}(S_i)}{k} \tag{22}$$

Consequently, the (k, T_B)-local Shannon entropy $\overline{H}(S)$ are used as the measure for describing the randomness over the entire test image S.

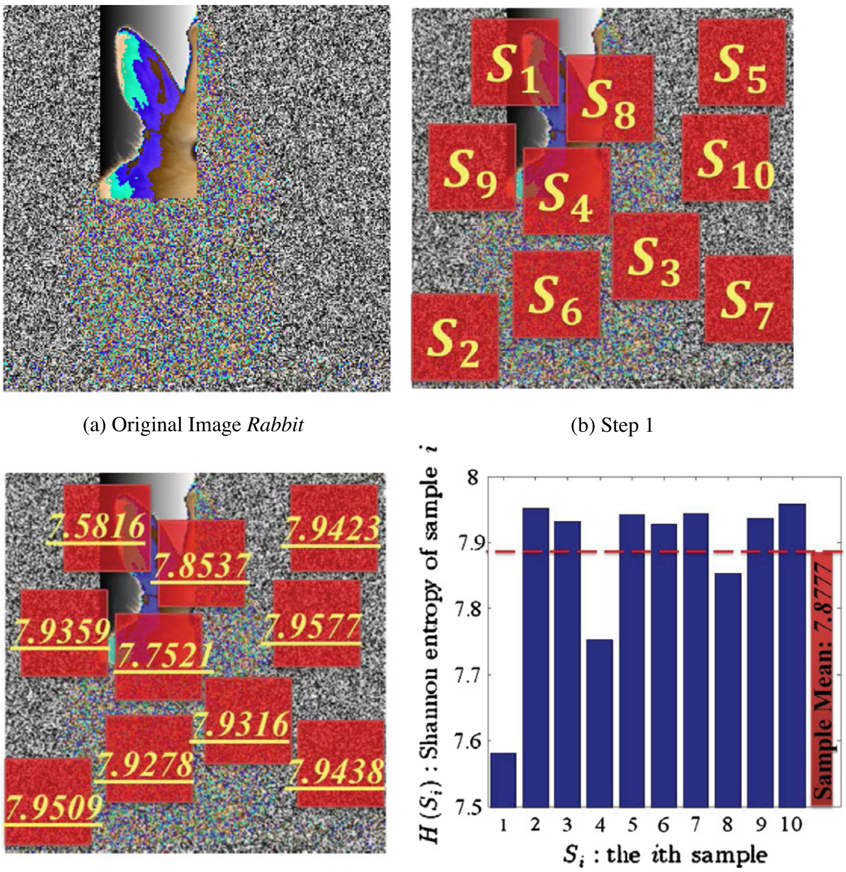
Fig. 3 shows the three steps of the (k, T_B)-local entropy measure on a poorly encrypted image, *Rabbit*, with $k = 9$ and $T_B = 256$. As can be seen in Fig. 3d, the local Shannon entropy directly points out the relatively low randomness scores for image blocks S_1, S_4, and S_8 in the test image, and thus demonstrates the capability of the measure to capture local randomness. Further, because the local image blocks are chosen randomly in the local Shannon entropy measure, randomness information of a test image is fairly represented in the resulting local Shannon entropy score.

Since this (k, T_B)-local Shannon entropy degrades to the global Shannon entropy for a T-pixel image S with $k = 1$ and $T_B = T$, the (k, T_B)-local Shannon entropy can be considered a generalized form of Shannon entropy, which includes the global Shannon entropy. However, this generalization is not trivial: one interpretation of the (k, T_B) parameter set is that this parameter set tunes the localization capacity of the Shannon entropy measure over a test image. A discussion on the parameters of the (k, T_B)-local Shannon entropy can be found in subsequent sections.

4.2. Why local and not global?

Because the local Shannon entropy measures image randomness by computing the sample mean of Shannon entropy over a number of non-overlapping and randomly selected image blocks, it is able to overcome some known weaknesses of the global Shannon entropy:

1. Inaccuracy: The global Shannon entropy sometimes fails to measure the true randomness of an image. Unlike global Shannon entropy, the (k, T_B) local Shannon entropy is able to capture local image block randomness a measure that might not be correctly reflected in the global Shannon entropy score.



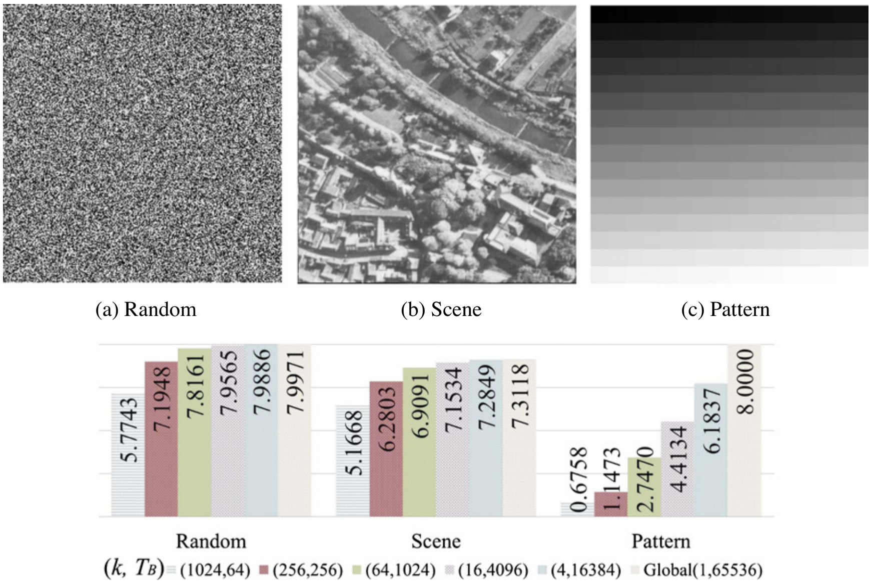
(a) Original Image *Rabbit*

(b) Step 1

(c) Step 2

(d) Step 3

Fig. 3. An example of the (k, T_B)-local entropy measure for k = 10, T_B = 256.



(a) Random

(b) Scene

(c) Pattern

(d) Averaged local Shannon entropy score

Fig. 4. Global and local Shannon entropy scores of sample images.

Table 2
Theoretical mean and standard deviation of (k, T_B)-local Shannon entropy score on random images.

$L = 2$	Binary image		$L = 256$	Grayscale image		$L = 256$	Color image	
T_B	$\mu_{\overline{H_{k,T_B}}(R)}$	$\sigma_{\overline{H_{k,T_B}}(R)}$	T_B	$\mu_{\overline{H_{k,T_B}}(R)}$	$\sigma_{\overline{H_{k,T_B}}(R)}$	T_B	$\mu_{\overline{H_{k,T_B}}(R)}$	$\sigma_{\overline{H_{k,T_B}}(R)}$
2×2	0.780639062	$0.307715375/\sqrt{k}$	2×2	1.988300234	$0.076064119/\sqrt{k}$	$2 \times 2 \times 3$	3.542339666	$0.082640020/\sqrt{k}$
4×4	0.953361607	$0.066187807/\sqrt{k}$	4×4	3.942064617	$0.082851351/\sqrt{k}$	$4 \times 4 \times 3$	5.407984610	$0.079041305/\sqrt{k}$
8×8	0.988638975	$0.016069236/\sqrt{k}$	8×8	5.765716929	$0.076603439/\sqrt{k}$	$8 \times 8 \times 3$	6.938975236	$0.059295884/\sqrt{k}$
16×16	0.997176704	$0.003992777/\sqrt{k}$	16×16	7.174966353	$0.052437999/\sqrt{k}$	$16 \times 16 \times 3$	7.737771412	$0.023253559/\sqrt{k}$
32×32	0.999295215	$0.000996718/\sqrt{k}$	32×32	7.808756571	$0.017246343/\sqrt{k}$	$32 \times 32 \times 3$	7.939203149	$0.005393141/\sqrt{k}$
64×64	0.999823868	$0.000249088/\sqrt{k}$	64×64	7.954588734	$0.004024888/\sqrt{k}$	$64 \times 64 \times 3$	7.984977322	$0.001330526/\sqrt{k}$

2. Inconsistency: The term ‘global’ is commonly inconsistent for images with various sizes, making the global Shannon entropy unsuitable as a universal measure. However, the (k, T_B)-local Shannon entropy is able to measure the image randomness using the same set of parameter regardless of the various sizes of test images and thus provides a relatively fair comparison for image randomness among multiple images.
3. Low efficiency: The global Shannon entropy measure requires the pixel information of an entire image, which is costly when the test image is large. However the local entropy measure requires only a portion of the total pixel information.

The first weakness regarding inaccuracy is illustrated in Fig. 4, which shows the global and local Shannon entropy scores for three classes of 8-bit grayscale images of size 256 × 256: *Random*, *Scene*, and *Pattern*. It should be noted that the global Shannon entropy score of image *Pattern* is exactly 8, which is the upper bound for an 8-bit grayscale image and implies image *Pattern* is very random-like, although image *Pattern* is not. In contrast, local Shannon entropy scores of image *Pattern* directly point out that this image is not random-like, because its entropy scores are obviously smaller than the corresponding scores of image *Random*. This example tells us that an image with a high global entropy score may not be necessarily random-like and that a random-like image always has high local entropy scores regardless of the used block size for the local measurement.

The second weakness regarding inconsistency is illustrated in Table 1, which shows that the expected Shannon entropy scores for different image sizes are very different. In general, an image of a larger size tends to have a higher Shannon entropy score than a smaller size image. Therefore, without considering its size, judging the randomness of an image from its global Shannon entropy score is pointless. In other words, measuring randomness using global Shannon entropy requires image size information whereas this is not the case with local Shannon entropy.

The low efficiency issue is straightforward and is addressed empirically in Section 6.

4.3. Hypothesis tests for the (k, T_B)-local Shannon entropy measure

As shown in Section 4.1 the (k, T_B)-local Shannon entropy is definitely a quantitative measure. In order to make it a qualitative measure, we derive hypothesis tests for the (k, T_B)-local Shannon entropy measure. This way, we can make a qualitative measure on a test image, namely reject or fail to reject a test image as a random image.

Consider the (k, T_B)-local Shannon entropy measure on a random image R as a random variable. According to the definition in Eq. (22), the following equations then hold

$$\mu_{\overline{H_{k,T_B}}(R)} = \mathrm{E}\left[\sum_{i=1}^{k} \frac{\mathrm{H}(R_i)}{k}\right] = \mu_{\mathrm{H}(X)} \tag{23}$$

$$\sigma^2_{\overline{H_{k,T_B}}(R)} = \mathrm{Var}\left[\sum_{i=1}^{k} \frac{\mathrm{H}(R_i)}{k}\right] = \frac{\sigma^2_{\mathrm{H}(X)}}{k} \tag{24}$$

where R_i denotes the ith image block in the random image R[1], and $\mu_{\mathrm{H}(X)}$ and $\sigma_{\mathrm{H}(X)}$ are the mean and standard deviation of the Shannon entropy score on a random image X, which has the same number of pixels in the random image block R_i.

Furthermore, this (k, T_B)-local Shannon entropy of a random image R actually follows a Gaussian distribution according to the CLT, as long as k is sufficiently large ($k \geqslant 30$). Namely,

$$\overline{H_{k,T_B}}(R) \sim \mathcal{N}(\mu_{\mathrm{H}(X)}, \sigma^2_{\mathrm{H}(X)}/k) \tag{25}$$

where X is a random image of the same format as R and also contains T_B pixels. Numerical results of (k, T_B)-local Shannon entropy with respect to various image sizes are given in Table 2.

Consequently, we design the (k, T_B)-local entropy hypothesis test for image randomness in the following way:

- Null hypothesis $\mathcal{H}_0$: $\overline{H_{k,T_B}}(S) = \mu_{\overline{H_{k,T_B}}(R)}$, which implies that the test image S is indistinguishable from a random image

[1] In the rest of paper, S and R always denotes a general test image and a random image for the (k,T_B)-local Shannon entropy measure, respectively.

- Alternative hypothesis $\mathcal{H}_1$: $\overline{H_{k,T_B}}(S) \neq \mu_{\overline{H_{k,T_B}}(R)}$, which implies that the test image S is distinguishable from a random image

Assuming $k \geqslant 30$ holds, the above test is a Z test, because the distribution of $\mathcal{H}_0$ is known (see Eq. (25)). As a result, the test statistic z can be obtained as follows

$$z = \frac{\overline{H_{k,T_B}}(S) - \mu_{\overline{H_{k,T_B}}(R)}}{\sigma_{\overline{H_{k,T_B}}(R)}} = \frac{\overline{H_{k,T_B}}(S) - \mu_{\mathrm{H}(X)}}{\sigma_{\mathrm{H}(X)}/\sqrt{k}} \tag{26}$$

With respect to the α-level of significance in a Z-test, we calculate critical values h^*_{left} and h^*_{right} using Eq. (27). Consequently, we fail to reject $\mathcal{H}_0$ for a test image S if $\overline{H_{k,T_B}}(S) \in \left[h^*_{left}, h^{l*}_{right}\right]$ otherwise we reject $\mathcal{H}_0$.

$$\begin{cases} h^*_{left} = \mu_{\overline{H_{k,T_B}}(R)} - \Phi^{-1}_{\alpha/2}\sigma_{\overline{H_{k,T_B}}(R)} = \mu_{\mathrm{H}(X)} - \Phi^{-1}_{\alpha/2}\sigma_{\mathrm{H}(X)}/\sqrt{k} \\ h^*_{right} = \mu_{\overline{H_{k,T_B}}(R)} + \Phi^{-1}_{\alpha/2}\sigma_{\overline{H_{k,T_B}}(R)} = \mu_{\mathrm{H}(X)} + \Phi^{-1}_{\alpha/2}\sigma_{\mathrm{H}(X)}/\sqrt{k} \end{cases} \tag{27}$$

where Φ^{-1} is the inverse cumulative density function of the standard normal distribution $\mathcal{N}(0,1)$.

4.4. Parameter selection

In the (k, T_B)-local entropy test, the two parameters play different roles. It is somewhat obvious that parameter k affects the width of the rejection region of the (k, T_B)-local entropy test. As k increases, $\mu_{\overline{H_{k,T_B}}(R)}$ remains the same, but $\sigma_{\overline{H_{k,T_B}}(R)}$ gets smaller, which implies a wider rejection region. However, since the (k, T_B)-local entropy test requires k to be sufficiently large to apply the CLT, k should be no less than 30. It should be noted that the condition $k \geqslant 30$ can be easily satisfied for most images.

The role of parameter T_B is more complex in the sense that both $\mu_{\overline{H_{k,T_B}}(R)}$ and $\sigma_{\overline{H_{k,T_B}}(R)}$ change as it changes. It is also noticeable that it is T_B that defines the local entropy test, because when T_B is as large as the number of the entire test image, then the local entropy test is identical to the global one. However, in contrast, if T_B is as small as one pixel, the local entropy test is pointless, as in this case $\overline{H_{k,T_B}}(S) = 0$, regardless of the test image S. Therefore, it is natural to ask whether there exists some *optimal* T_B in between these two extremes.

Since there are multiple ways to address the problem, we made the choice to approach it by first considering the two parameters below:

- C_{local} (Capacity to capture local randomness information): This capacity is reduced as T_B increases. When T_B increases to the number of pixels in the entire test image, the (k, T_B)-local entropy measure becomes the exact global measurement, and thus does not capture any local randomness information.
- C_{scale} (Capacity to capture all L levels of intensity scales): This capacity is enhanced as T_B increases. Although a local entropy score is supposed to measure local randomness with T_B pixels and L intensity scales, such a measure is inaccurate when $T_B \ll L$, because the number of possible symbols that can be observed within T_B is always less than or equal to min (T_B, L). When T_B = 1, the (k, T_B)-local entropy measure is always 0, and thus completely loses this capacity.

C_{local} and C_{scale} can be modeled as follows, and consequently we can construct a capacity energy function f as shown as follows:

$$\begin{aligned} &C_{local}(T_B) = 1/T_B \\ &C_{scale}(T_B) = \Pr(\text{see all } L \text{ scales in } T_B \text{ pixels}|L) \\ &f^L(T_B) = C_{local}(T_B)\cdot C_{scale}(T_B) = \Pr\ (\text{see all } L \text{ scales in } T_B \text{ pixels}|L)/T_B \end{aligned} \tag{28}$$

This energy function can be interpreted as the capacity to capture all L scales per pixel.

It is worthwhile to note that Pr (see all L scales in $T_B$pixels|L) can be calculated via Eq. (29) [26], where the random variable X = the number of drawings it takes to get all L intensity scales in Pr (X = t|L).

$$\Pr(\text{see all } L \text{ scales in } T_B \text{ pixels}|L) = \sum_{t=1}^{T_B}\Pr(X = t|L) \tag{29}$$

$$\Pr(X = t|L) = \frac{1}{L^t}\sum_{i=1}^{L}(-1)^{i-1}\cdot i\cdot\binom{L}{i}\cdot(L-i)^{t-1} \tag{30}$$

When $T_B \to 1$, C_{scale} decreases to 0 and thus $f^L(T_B)$ quickly goes to 0; when $T_B \to \infty$, C_{local} goes to infinity and thus $f^L(T_B)$ also approaches 0. As a result, the optimized T_B^{L*} with respect to L intensity scales can be obtained as follows

$$T_B^{L*} = \arg\max_{T_B \geqslant 1} f^L(T_B) \tag{31}$$

Fig. 5 shows $f^L(T_B)$ scores with respect to different L values. As expected, $f^L(T_B)$ is a unimodal function. More specifically,

$$T_B^{L=2*} = 2 \tag{32}$$

$$T_B^{L=256*} = 1936 \tag{33}$$

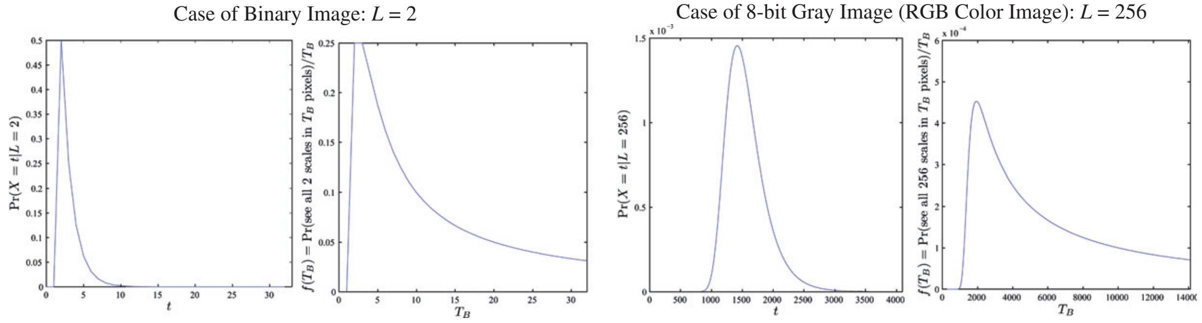


Fig. 5. T_B optimization with respect to different image formats (left column: C_{scale}; right column: capacity energy).

Table 3
Mean and standard deviation of local Shannon entropy for random images with optimal T_B.

Image type	L	T_B^{L*}	$\mu_{\mathrm{H}(X)}$	$\sigma_{\mathrm{H}(X)}$	$\mu_{\overline{\mathrm{H}_{k,T_B}}(R)}$	$\sigma_{\overline{\mathrm{H}_{k,T_B}}(R)}$
Binary	2	2	0.500000000	0.500000000	0.500000000	$0.500000000/\sqrt{k}$
Grayscale (8-bit)	256	1936	7.902469317	0.008694225	7.902469317	$0.008694225/\sqrt{k}$
Color (RGB)	256	1936	7.902469317	0.008694225	7.902469317	$0.008694225/\sqrt{k}$

Table 4
Local entropy test reference table for optimized T_B.

k	α = 0.05		α = 0.01		α = 0.001	
	h_{left}^{l*}	h_{right}^{l*}	h_{left}^{l*}	h_{right}^{l*}	h_{left}^{l*}	h_{right}^{l*}
Binary image $L = 2,\ T_B^{L=2*} = 2$						
30	0.467333934	0.532666066	0.457069512	0.542930488	0.445157888	0.554842112
40	0.475500450	0.524499550	0.467802134	0.532197866	0.458868416	0.541131584
50	0.480400360	0.519599640	0.474241707	0.525758293	0.467094733	0.532905267
60	0.483666967	0.516333033	0.478534756	0.521465244	0.472578944	0.527421056
70	0.486000257	0.513999743	0.481601219	0.518398781	0.476496238	0.523503762
80	0.487750225	0.512249775	0.483901067	0.516098933	0.479434208	0.520565792
90	0.489111311	0.510888689	0.485689837	0.514310163	0.481719296	0.518280704
100	0.490200180	0.509799820	0.487120853	0.512879147	0.483547366	0.516452634
110	0.491091073	0.508908927	0.488291685	0.511708315	0.485043060	0.514956940
120	0.491833483	0.508166517	0.489267378	0.510732622	0.486289472	0.513710528
130	0.492461677	0.507538323	0.490092964	0.509907036	0.487344128	0.512655872
140	0.493000129	0.506999871	0.490800610	0.509199390	0.488248119	0.511751881
150	0.493466787	0.506533213	0.491413902	0.508586098	0.489031578	0.510968422
8-bit Grayscale/RBG color image $L = 2,\ T_B^{L=256*} = 1936$						
30	7.901901305	7.903037329	7.901722822	7.903215812	7.901515698	7.903422936
40	7.902043308	7.902895326	7.901909446	7.903029188	7.901754103	7.903184531
50	7.902128510	7.902810124	7.902021420	7.902917214	7.901897145	7.903041489
60	7.902185311	7.902753323	7.902096070	7.902842564	7.901992507	7.902946127
70	7.902225883	7.902712751	7.902149391	7.902789243	7.902060623	7.902878011
80	7.902256312	7.902682322	7.902189382	7.902749252	7.902111710	7.902826924
90	7.902279980	7.902658654	7.902220485	7.902718149	7.902151444	7.902787190
100	7.902298913	7.902639721	7.902245369	7.902693265	7.902183231	7.902755403
110	7.902314405	7.902624229	7.902265728	7.902672906	7.902209239	7.902729395
120	7.902327314	7.902611320	7.902282693	7.902655941	7.902230912	7.902707722
130	7.902338237	7.902600397	7.902297049	7.902641585	7.902249251	7.902689383
140	7.902347600	7.902591034	7.902309354	7.902629280	7.902264970	7.902673664
150	7.902355715	7.902582919	7.902320018	7.902618616	7.902278593	7.902660041

Table 3 gives the required statistics for local entropy tests with respect to image types and Table 4 provides acceptance intervals of (k, T_B)-local entropy tests under various significance levels with respect to parameter k.

5. Simulation results

In this section, we compare the theoretical distributions and statistics derived in previous sections to those observed from a random image source. Specifically speaking,

- Theoretical statistics of a random image in Table 1 vs. observed ones
- Theoretical distribution of the (k, T_B)-local entropy score for random images vs. observed distribution

All simulations are performed under the environment of MATLAB r2011a using built-in functions.

5.1. Estimated statistics from observed random images

Table 5 shows the mean and standard deviations of 50,000 observed sample images $S^{(1)}, S^{(2)}, \ldots, S^{(50{,}000)}$ generated from the MATLAB uniform RNG, where $S^{(i)}$ denotes the ith sample out of 50,000. The sample mean and the sample standard deviation for each parameter setting are calculated from 50,000 observations using Eqs. (34) and (35), respectively. It is noticeable that estimated values are very close to theoretical values in Table 1

$$\widehat{\mu_{\mathrm{H}(S)}} = \sum_{i=1}^{50{,}000} \mathrm{H}(S^{(i)})/50,000 \tag{34}$$

$$\widehat{\sigma_{\mathrm{H}(S)}} = \sqrt{\left(\sum_{i=1}^{50{,}000} \mathrm{H}(S_i)\right)^2 - 50,000(\widehat{\mu_{\mathrm{H}(S)}})^2/49,999} \tag{35}$$

Table 6 tabulates a quantization of the errors between Tables 1 and 5 using *mean square error* (MSE) as shown in Eq. (36). Symbol $\hat{\theta}$ is the estimated value of the true parameter θ. In our case, $\hat{\theta}$ is $\widehat{\mu_{\mathrm{H}(S)}}$ or $\widehat{\sigma_{\mathrm{H}(S)}}$ and θ is the corresponding $\mu_{\mathrm{H}(X)}$ or $\sigma_{\mathrm{H}(X)}$ under the same image settings in Table 1. As seen in this table, the observed statistics from a random image source match very well with those we derived.

$$\mathrm{MSE}(\hat{\theta}) = \mathrm{E}\left((\hat{\theta} - \theta)^2\right) \tag{36}$$

5.2. Observed distribution of z statistic in the (k, T_B)-local entropy test

We showed earlier that the test statistic $z = \frac{\overline{H_{k,T_B}}(S) - \mu_{\overline{H_{k,T_B}}(R)}}{\sigma_{\overline{H_{k,T_B}}(R)}}$ follows the standard Normal distribution, i.e. $z \sim \mathcal{N}(0, 1)$, where $\overline{H_{k,T_B}}(S)$ and $\overline{H_{k,T_B}}(R)$ are (k,T_B)-local entropy scores for a test image S and a random image R (see Table 2 for details).

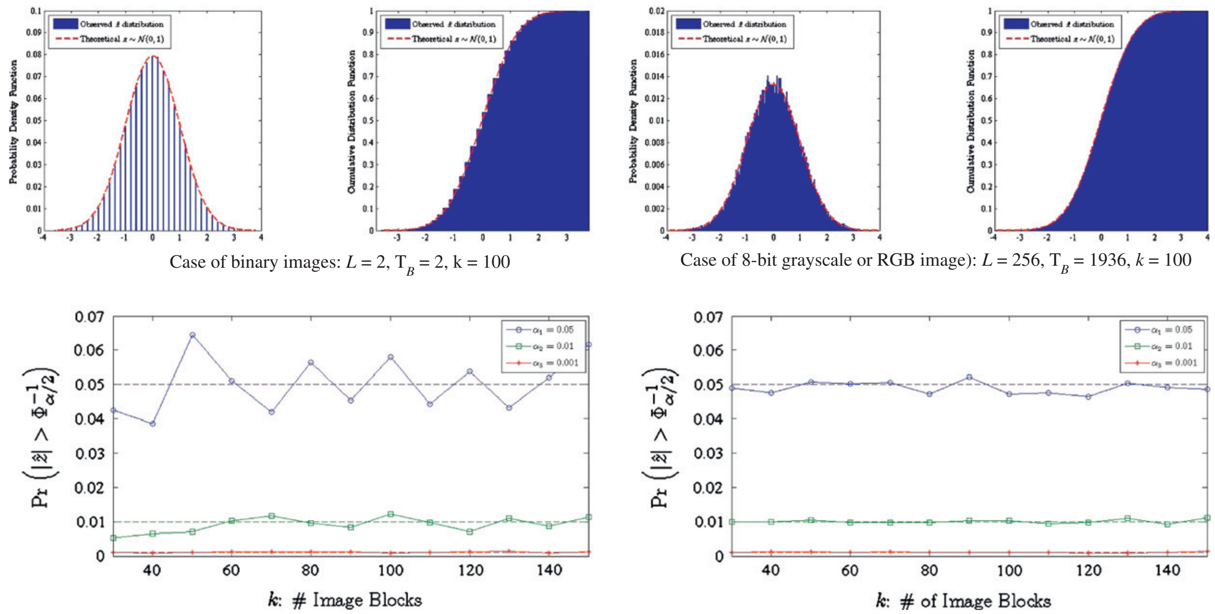
We construct an observed distribution about $\hat{z}$ which is the normalized histogram of $\frac{\overline{H_{k,T_B}}(S^{(i)}) - \mu_{\overline{H_{k,T_B}}(R)}}{\sigma_{\overline{H_{k,T_B}}(R)}}$ using 256 bins, where $S^{(i)}$ denotes the ith observed images from the MATLAB RNG and $1 \leqslant i \leqslant 50{,}000$. Fig. 6 shows the estimated distribution $\hat{z}$ and

Table 5
Observed mean and standard deviation of Shannon entropy scores for random images.

$L = 2$	Binary image		$L = 256$	Grayscale image		$L = 256$	Color image	
T	$\widehat{\mu_{\mathbf{H}(S)}}$	$\widehat{\sigma_{\mathbf{H}(S)}}$	T	$\widehat{\mu_{\mathbf{H}(S)}}$	$\widehat{\sigma_{\mathbf{H}(S)}}$	T	$\widehat{\mu_{\mathbf{H}(S)}}$	$\widehat{\sigma_{\mathbf{H}(S)}}$
2 × 2	0.779166957	0.308881174	2 × 2	1.988628677	0.075070562	2 × 2 × 3	3.541664155	0.083632601
4 × 4	0.953636046	0.065588142	4 × 4	3.942070485	0.082578193	4 × 4 × 3	5.408002228	0.079179351
8 × 8	0.988651881	0.016031081	8 × 8	5.765586044	0.076379599	8 × 8 × 3	6.938676283	0.058889500
16 × 16	0.997184215	0.003984394	16 × 16	7.174563315	0.052355799	16 × 16 × 3	7.737746839	0.023312601
32 × 32	0.999289182	0.001007253	32 × 32	7.808835357	0.017273323	32 × 32 × 3	7.939155968	0.005428824
64 × 64	0.999823294	0.000250619	64 × 64	7.954594365	0.004022400	64 × 64 × 3	7.984975098	0.001335346

Table 6
Mean square error between theoretical statistics in Table 1 and observed statistics in Table 5.

$L = 2$	Binary image		$L = 256$	Grayscale image		$L = 256$	Color image	
T	$\mathrm{MSE}(\widehat{\mu_{\mathbf{H}(S)}})$	$\mathrm{MSE}(\widehat{\sigma_{\mathbf{H}(S)}})$	T	$\mathrm{MSE}(\widehat{\mu_{\mathbf{H}(S)}})$	$\mathrm{MSE}(\widehat{\sigma_{\mathbf{H}(S)}})$	T	$\mathrm{MSE}(\widehat{\mu_{\mathbf{H}(S)}})$	$\mathrm{MSE}(\widehat{\sigma_{\mathbf{H}(S)}})$
2 × 2	2.16709E−006	1.35909E−006	2 × 2	1.07874E−007	9.87154E−007	2 × 2 × 3	4.56316E−007	9.85216E−007
4 × 4	7.53167E−008	3.59597E−007	4 × 4	3.44258E−011	7.46151E−008	4 × 4 × 3	3.10404E−010	1.90566E−008
8 × 8	1.66555E−010	1.45580E−009	8 × 8	1.71308E−008	5.01041E−008	8 × 8 × 3	8.93732E−008	1.65148E−007
16 × 16	5.64146E−011	7.02887E−011	16 × 16	1.62439E−007	6.75684E−009	16 × 16 × 3	6.03824E−010	3.48602E−009
32 × 32	3.63932E−011	1.10991E−010	32 × 32	6.20723E−009	7.27957E−010	32 × 32 × 3	2.22599E−009	1.27331E−009
64 × 64	3.29605E−013	2.34526E−012	64 × 64	3.17103E−011	6.19361E−012	64 × 64 × 3	4.94678E−012	2.32330E−011



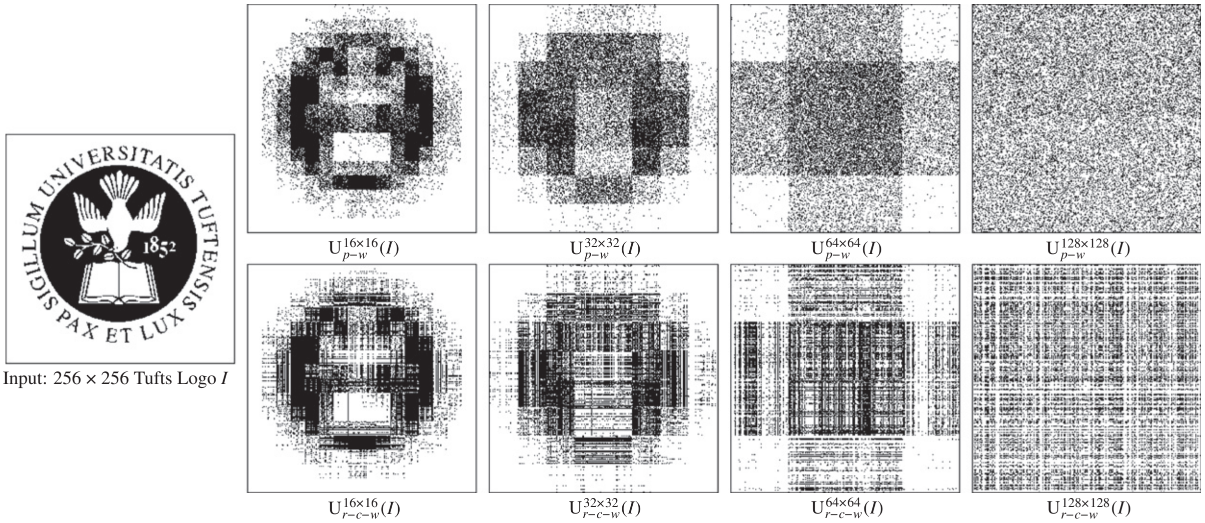
Case of binary images: $L = 2$, $T_B = 2$, k = 100

Case of 8-bit grayscale or RGB image): $L = 256$, $T_B = 1936$, $k = 100$

Sample percentage out of α confidence interval $L = 2$, $T_B = 2$

Sample percentage out of α confidence interval $L = 256$, $T_B = 1936$

Fig. 6. Observed distribution of (k, T_B)-local entropy test statistic vs. theoretical distribution $z \sim \mathcal{N}(0, 1)$.



Input: 256×256 Tufts Logo I

$\mathrm{U}^{16\times16}_{p-w}(I)$ $\mathrm{U}^{32\times32}_{p-w}(I)$ $\mathrm{U}^{64\times64}_{p-w}(I)$ $\mathrm{U}^{128\times128}_{p-w}(I)$

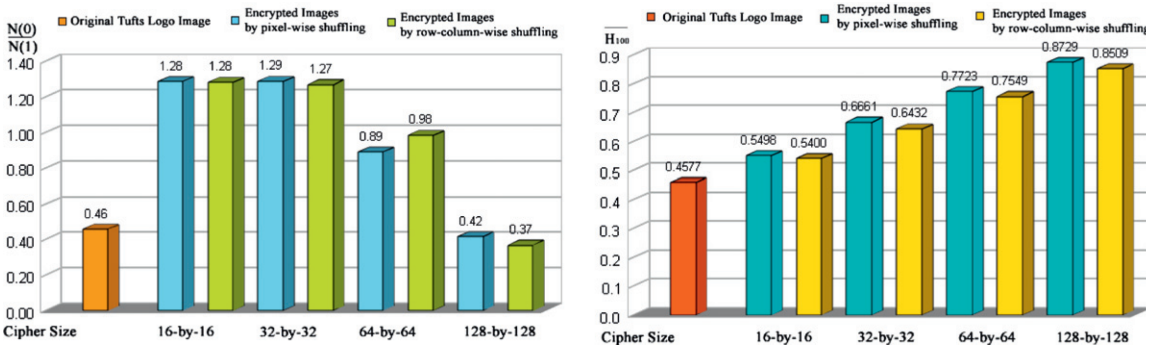
$\mathrm{U}^{16\times16}_{r-c-w}(I)$ $\mathrm{U}^{32\times32}_{r-c-w}(I)$ $\mathrm{U}^{64\times64}_{r-c-w}(I)$ $\mathrm{U}^{128\times128}_{r-c-w}(I)$

Fig. 7. Image shuffling results.

the theoretical distribution of z for the (k, T_B)-local entropy test, where the first row shows the observed $\hat{z}$ vs. the standard normal distribution, with respect to the optimized $T_B = 2$ for binary images and $T_B = 1936$ for grayscale or RGB images for a fixed number of samples $k = 100$; and the second row shows the sample percentage out of the α-level confidence interval for $k = 30, 40, \ldots, 150$, with respect to $\alpha = 0.05, 0.01, 0.001$ (see Table 4). As seen in Fig. 6, the observed distribution $\hat{z}$ is very close to the standard Normal distribution, which is a conclusion discussed earlier; and the actual distribution tails, i.e. $\Pr\left(|\hat{z}| > \Phi^{-1}_{\alpha/2}\right)$ are close to their theoretical values. Additionally, as the number of image blocks k in the (k, T_B)-local entropy test increases, $\Pr\left(|\hat{z}| > \Phi^{-1}_{\alpha/2}\right)$ gets closer to zero.

Table 7
Pixel randomness for shuffled images in Fig. 7.

Images	FIPS 140-2 tests										Local entropy score		Global entropy score
	Monobit		Poker	Run						Long run			
				Length of the run									
				1	2	3	4	5	⩾6		Mean	Std	
I	*0 bit*	11299	2913.984	106	131	109	36	42	280	110	0.4576711	0.3697290	0.95995162
	1 bit	8701		53	70	59	54	73	395				
$U_{p-w}^{16\times16}(I)$	*0 bit*	11244	2386.291	1371	530	292	137	101	285	82	0.5498039	0.3324658	0.9599516
	1 bit	8756		1238	565	309	169	84	352				
$U_{p-w}^{32\times32}(I)$	*0 bit*	11251	695.8784	1774	830	499	239	184	461	7	0.6660959	0.3168008	0.9599516
	1 bit	8749		2232	828	370	199	113	246				
$U_{p-w}^{64\times64}(I)$	*0 bit*	9411	869.2224	2354	1050	512	263	155	218	0	0.7723314	0.2896868	0.9599516
	1 bit	10589		2105	1112	550	298	165	323				
$U_{p-w}^{128\times128}(I)$	*0 bit*	5872	4582.874	2897	856	259	79	26	6	0	0.8728581	0.0552799	0.9599516
	1 bit	14128		1174	862	592	458	315	722				
$U_{r-c-w}^{16\times16}(I)$	*0 bit*	11235	2802.157	1137	348	146	114	26	222	97	0.5399521	0.3266322	0.9599516
	1 bit	8765		857	481	190	122	43	301				
$U_{r-c-w}^{32\times32}(I)$	*0 bit*	11179	822.0224	1457	476	212	138	67	339	83	0.6431998	0.3261251	0.9599516
	1 bit	8821		1259	610	157	190	119	355				
$U_{r-c-w}^{64\times64}(I)$	*0 bit*	9917	377.1584	1535	600	302	115	112	348	57	0.7548611	0.3052360	0.9599516
	1 bit	10083		1476	574	338	109	139	375				
$U_{r-c-w}^{128\times128}(I)$	*0 bit*	5370	10966.144	1957	714	384	79	56	34	0	0.8508929	0.1121833	0.9599516
	1 bit	14630		1075	752	449	269	156	524				



(a)Tendencyof $T = \#0bits\ /\ \#1bits$

(b) Tendency of $\overline{H_{100,256}}(S)$

Fig. 8. Randomness test for image shuffling.

6. Applications to image encryption

In this section, we demonstrate two applications of the (k, T_B)-local Shannon entropy measure, where the first one uses the local Shannon entropy measure to evaluate the image shuffling performance whereas the global entropy measure is ineffective, and the second one uses the derived (k, T_B)-local Shannon entropy test to evaluate the performance of various image ciphers and compares computational costs of the global and local Shannon entropy measures.

6.1. Evaluating pixel randomness for image shuffling

Generally speaking, a pixel (-wise) shuffling algorithm for image encryption can be defined as in Eq. (37), where I and O denote input and output images respectively; $O(i,j)$ indicates the pixel located at intersection of the ith row and the jth column in the output image O; $U(\cdot)$ is a general image shuffling algorithm; e_{Π}^{rc} is a pixel permutation. Additionally, if Eq. (38) holds, that is if e_{Π}^{rc} is the combination of a row permutation e_{Π}^{r} and a column permutation e_{Π}^{c}, then a pixel (-wise) shuffling is equivalent to a row-column (-wise) shuffling. As its definition implies, a shuffling algorithm does not change the intensity level for any pixel, but shuffles the positions of pixels.

$$O(i,j) = U(I) = I(e_{\Pi}^{rc}(i,j)) \qquad (37)$$

$$I(e_{\Pi}^{rc}(i,j)) = I(e_{\Pi}^{r}(i), e_{\Pi}^{c}(j)) \qquad (38)$$

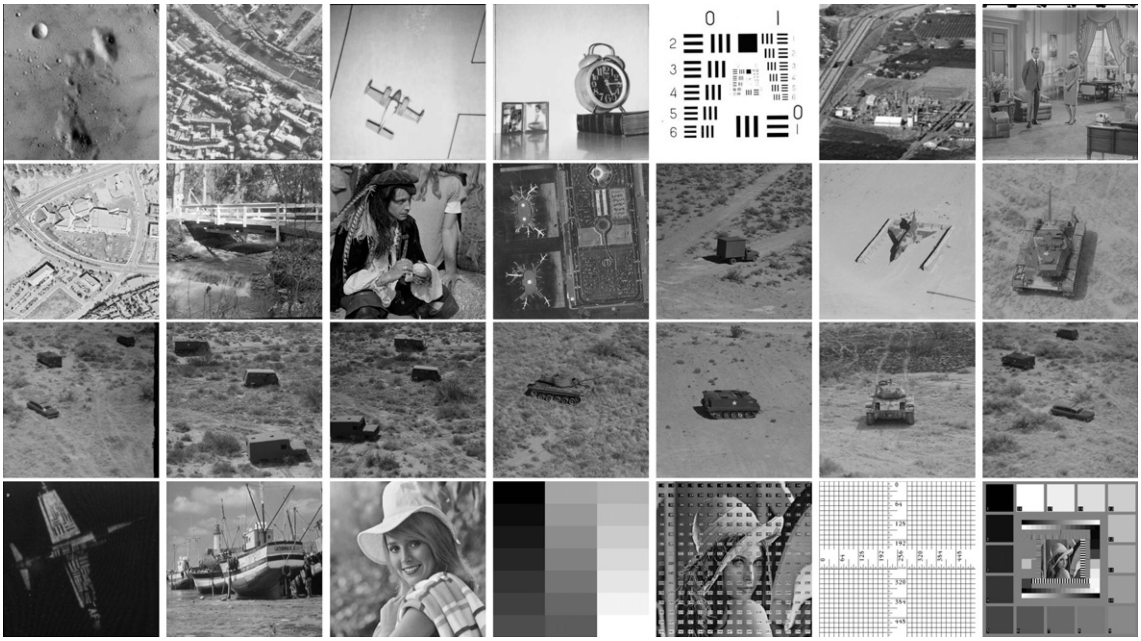


Fig. 9. Selected images in USC-SIPI *Miscellaneous* dataset.

Table 8
Global Shannon entropy for encrypted images.

File	Image information			Global Shannon entropy of encrypted images			
	Description	Size	Type	bmpPacker	I-Cipher	3DCat	Sudoku
5.1.09	Moon surface	256 × 256	Gray	7.8581	7.9990	7.9972	7.9972
5.1.10	Aerial	256 × 256	Gray	7.8581	7.9991	7.9975	7.9970
5.1.11	Airplane	256 × 256	Gray	7.8586	7.9991	7.9972	7.9974
5.1.12	Clock	256 × 256	Gray	7.8574	7.9992	7.9970	7.9974
5.1.13	Resolution chart	256 × 256	Gray	6.1780	7.9990	7.9928	7.9947
5.1.14	Chemical plant	256 × 256	Gray	7.8562	7.9990	7.9973	7.9970
5.2.08	Couple	512 × 512	Gray	7.9911	7.9998	7.9991	7.9993
5.2.09	Aerial	512 × 512	Gray	7.9915	7.9998	7.9993	7.9993
5.2.10	Stream and bridge	512 × 512	Gray	7.9912	7.9998	7.9993	7.9992
5.3.01	Man	1024 × 1024	Gray	7.9993	8.0000	7.9996	7.9998
5.3.02	Airport	1024 × 1024	Gray	7.9993	7.9999	7.9998	7.9998
7.1.01	Truck	512 × 512	Gray	7.9916	7.9997	7.9993	7.9992
7.1.02	Airplane	512 × 512	Gray	7.9915	7.9998	7.9993	7.9990
7.1.03	Tank	512 × 512	Gray	7.9914	7.9998	7.9993	7.9992
7.1.04	Car and APCs	512 × 512	Gray	7.9917	7.9998	7.9993	7.9992
7.1.05	Truck and APCs	512 × 512	Gray	7.9910	7.9997	7.9993	7.9992
7.1.06	Truck and APCs	512 × 512	Gray	7.9921	7.9997	7.9992	7.9993
7.1.07	Tank	512 × 512	Gray	7.9914	7.9998	7.9993	7.9992
7.1.08	APC	512 × 512	Gray	7.9915	7.9998	7.9991	7.9991
7.1.09	Tank	512 × 512	Gray	7.9915	7.9998	7.9993	7.9992
7.1.10	Car and APCs	512 × 512	Gray	7.9917	7.9997	7.9993	7.9992
7.2.01	Airplane (U-2)	1024 × 1024	Gray	7.9993	7.9999	7.9998	7.9997
Boat.512	Fishing boat	512 × 512	Gray	7.9914	7.9997	7.9993	7.9993
Elaine.512	Girl (Elaine)	512 × 512	Gray	7.9914	7.9998	7.9993	7.9992
Gray21.512	21 level step wedge	512 × 512	Gray	7.5291	7.9998	7.9992	7.9997
Numbers.512	256 level test pattern	512 × 512	Gray	7.9916	7.9997	7.9994	7.9993
Ruler.512	Pixel ruler	512 × 512	Gray	5.2343	7.9998	6.9995	7.9943
Testpat.1k	General test pattern	1024 × 1024	Gray	7.5425	7.9999	7.9820	7.9995
			Mean	7.7726	7.9996	7.9624	7.9986
			Std	0.6102	0.0003	0.1887	0.0014

A shuffling algorithm can also be a block cipher, which shuffles an image block by block.

Fig. 7 shows image shuffling results using the pixel (-wise) shuffling and row-column (-wise) shuffling. For simplicity, consider a shuffled image I using a pixel (-wise) shuffling scheme with an M-by-N block size:

Table 9
Local Shannon entropy test for encrypted images ($k = 30$, $T_B = 1936$, $\alpha = 0.05$).

Filename	Local Shannon entropy of encrypted images							
	bmpPacker		I-Cipher		3DCat		Sudoku	
	Score	*P* value	Score	*P* value	Score	*P* value	Score	*P* value
5.1.09	7.7616	0.0000	7.8994	0.0519	7.9032	0.6334	7.9035	0.5305
5.1.10	7.7640	0.0000	7.9009	0.3168	7.9021	0.8302	7.9042	0.2774
5.1.11	7.7673	0.0000	7.9032	0.6543	7.9034	0.5564	7.9092	0.0000
5.1.12	7.7619	0.0000	7.9046	0.1755	7.9038	0.3943	7.9075	0.0015
5.1.13	6.0686	0.0000	7.8996	0.0726	7.7877	0.0000	7.9240	0.0000
5.1.14	7.7593	0.0000	7.9003	0.1626	7.9012	0.4073	7.8985	0.0116
5.2.08	7.8954	0.0000	7.9002	0.1483	7.9012	0.4221	7.9026	0.9401
5.2.09	7.8940	0.0000	7.9011	0.4002	7.9043	0.2440	7.9049	0.1249
5.2.10	7.8947	0.0000	7.9041	0.3084	7.9015	0.5336	7.9028	0.8142
5.3.01	7.9027	0.8823	7.9023	0.9170	7.8226	0.0000	7.9041	0.3136
5.3.02	7.9048	0.1437	7.9031	0.7074	7.8714	0.0000	7.8978	0.0034
7.1.01	7.8947	0.0000	7.9015	0.5299	7.9012	0.4300	7.9014	0.4809
7.1.02	7.8957	0.0000	7.9050	0.1170	7.8962	0.0001	7.9036	0.4741
7.1.03	7.8962	0.0001	7.9049	0.1303	7.9032	0.6664	7.9019	0.7432
7.1.04	7.8957	0.0000	7.9025	0.9830	7.9009	0.3183	7.8991	0.0337
7.1.05	7.8905	0.0000	7.9012	0.4144	7.9033	0.6161	7.8993	0.0477
7.1.06	7.8944	0.0000	7.9021	0.8397	7.9053	0.0761	7.9039	0.3596
7.1.07	7.8972	0.0009	7.9071	0.0035	7.9030	0.7303	7.9024	0.9406
7.1.08	7.8928	0.0000	7.9016	0.5652	7.8945	0.0000	7.9003	0.1780
7.1.09	7.8963	0.0001	7.9040	0.3373	7.9040	0.3248	7.9000	0.1164
7.1.10	7.8958	0.0000	7.9021	0.8344	7.9017	0.6232	7.9015	0.5312
7.2.01	7.9023	0.9143	7.9031	0.7046	7.8690	0.0000	7.8976	0.0023
Boat.512	7.8927	0.0000	7.9025	0.9644	7.9004	0.1921	7.9051	0.0975
Elaine.512	7.8962	0.0001	7.9003	0.1663	7.9029	0.7968	7.9027	0.9015
Gray21.512	7.4038	0.0000	7.9026	0.9141	7.8895	0.0000	7.9202	0.0000
Numbers.512	7.8963	0.0001	7.9019	0.7185	7.8976	0.0020	7.9020	0.7679
Ruler.512	4.9873	0.0000	7.9020	0.7794	6.8212	0.0000	7.8243	0.0000
Testpat.1 k	6.6193	0.0000	7.9031	0.6875	7.7283	0.0000	7.8668	0.0000
Mean ± Std	7.6401 ± 0.6646		7.9024 ± 0.0018		7.8473 ± 0.2052		7.8997 ± 0.0174	
# Images passed α-level test	3		27		18		17	

- $U^{M\times N}(I)$ gets more random-like as the block size $M \times N$ increases
- $U^{M\times N}_{p-w}(I)$ is more random-like than $U^{M\times N}_{r-c-w}(I)$

Table 7 shows the image pixel randomness measured by the FIPS 140-2 tests [4], local Shannon entropy and global Shannon entropy for shuffled images in Fig. 7. The global Shannon entropy fails to differentiate image randomness under various shuffling methods. This is because global Shannon entropy measures randomness over the entire image, which remains the same regardless of shuffling methods. The local Shannon entropy measures image randomness with 100 local image blocks each with 256 pixels, i.e. $k = 100$, $T_B = 256$. The FIPS 140-2 tests require a bit-string of length 20,000, which is given as the first 20,000 pixels in a shuffled image. It is important to note that the local Shannon entropy scores give the same conclusions about the pixel randomness of shuffled images in Fig. 7 as the conclusions derived earlier, from a human visual perspective. Although similar conclusions can be drawn using the ratios of *0-bit* to *1-bit* in the Monobit test of FIPS 140-2 test suite, it is an indirect approach, and not as salient as the results from the local Shannon entropy measure. Fig. 8 compares the trends given by the Monobit test and the local Shannon entropy measure.

6.2. Evaluating image randomness for image encryption

In this section, the local Shannon entropy measure is applied to encrypted images from four image ciphers: commercial image ciphers *I-Cipher*[2], *bmpPacker*[3], and image encryption algorithms *3DCat* [16] and *Sudoku* [60]. It is worthwhile to note that the reason why the local Shannon entropy test is able to differentiate an encrypted image generated by a secure cipher from those generated by a insecure cipher is that a secure image cipher attains the confusion and diffusion properties [48], which implies that it could be a pseudo RIG. Therefore, the proposed local Shannon entropy test for random images is applicable to encrypted images.

We first selected 28 original images from the USC-SIPI image database[4] (see Fig. 9) and then encrypted these images using the four image ciphers/algorithms. Finally, we applied the global and *(30,1936)*-local Shannon entropy measures to all encrypted images and obtained Tables 8 and 9.

[2] *I-Cipher* is a product of *Ambitware Inc.*, available at http://www.ambitware.com/ under *product* page as the date of 07/19/2012.

[3] *bmpPacker* is written by Jens Ğodeke, available at http://www.jens-goedeke.eu/tools/bmppacker/ as the date of 07/19/2012.

[4] A public image database distributed by the University of Southern California, available at http://sipi.usc.edu/database/ as the date of 07/19/2012.

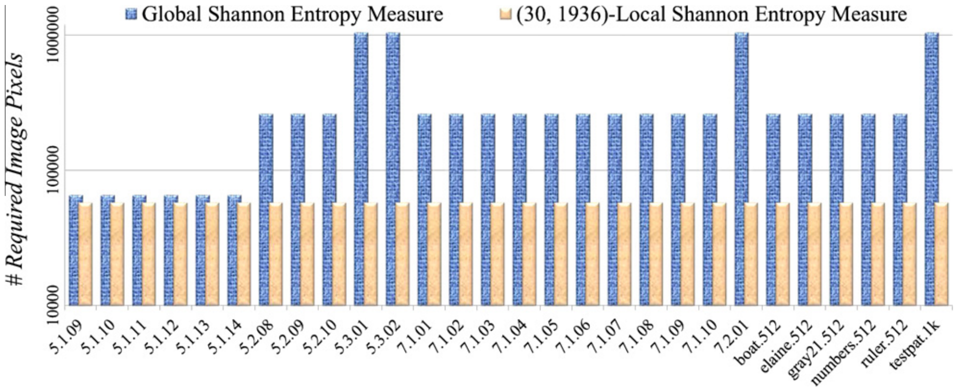


Fig. 10. Complexity comparisons between the local and the global Shannon entropy measure.

Comparing Tables 8 and 9, it can be observed that although the performance of the four ciphers in the global and local Shannon entropy measures gives the same ranking, *I-Cipher* > *Sudoku* > *3DCat* > *bmpPacker*, the global Shannon entropy measure shows the performance of *Sudoku* to be much closer to that of *I-Cipher* than to that of *3DCat*, while the local Shannon entropy test indicates that this conclusion may be wrong, in the sense that *3DCat* has an even better acceptance rate than *Sudoku*. This shows the importance of the qualitative analysis provided by the local Shannon entropy test. Finally, we compare the number of required image pixels for the global and local Shannon entropy tests in Fig. 10, which clearly shows that the global Shannon entropy measure requires a different number of pixels for different images, but always much more than those required by the local Shannon entropy measure.

7. Conclusions

In this paper, we have introduced the concept of the (k, T_B)-local Shannon entropy measure for image randomness, which includes the global Shannon entropy as a special case. The proposed $(k, T_B) -$ local Shannon entropy measure is defined on k local image blocks with T_B pixels and computes the sample mean of the Shannon entropy in each image block. Therefore, the proposed local Shannon entropy measure.

1. is able to measure image randomness by quantizing the captured local randomness information;
2. requires a fewer number of image pixels to compute an entropy score and thus is faster than the global Shannon entropy measure;
3. allows fair randomness comparisons between images of different sizes.

Furthermore, we have derived the hypothesis tests of the (k, T_B)-local Shannon entropy measure for random images, providing a qualitative measure to a test image. Consequently, the (k, T_B)-local Shannon entropy measure can be directly used to test whether a test image is random-like by simply comparing an observed score to a theoretical one.

Our simulation results have demonstrated that the estimated statistics and observed distributions of the local Shannon entropy from 50,000 random image samples fit our proposed mathematical model very well. Finally, we showed possible applications of the local Shannon entropy measure in image shuffling and image encryption. Our tested results showed that the proposed local Shannon entropy is a better measure for image shuffling than those available in FIPS 140-2, and that the local Shannon entropy scores on various shuffled images also reflect human intuition. Moreover, the test results of local Shannon entropy scores imply that many image ciphers/algorithms are not as random-like as claimed by their authors. Three out of four tested image ciphers actually have much lower local Shannon entropy scores than expected for securely encrypted images.

One possible remedy to avoid producing these poorly encrypted images is to use the local Shannon entropy test as a quality control device in a feedback encryption system: if an encrypted image fails to pass the local Shannon entropy test, then it should be sent back to the cipher and be encrypted again. In this manner, the local Shannon entropy test would ensure that all encrypted images processed by an image cipher are random-like and indistinguishable from random images.

Acknowledgements

Yue Wu is sponsored by Raytheon BBN Technologies in Cambridge Massachusetts. We would like to thank Professor Marjorie Hahn of the Tufts University Mathematics Department for her suggestions and help on this paper, and Jincheng Pang of the Tufts University Electrical and Computer Engineering Department for his suggestions and long time support.

References

[1] FIPS PUB 140-1: Security Requirements for Cryptographic Modules, 1994.
[2] FIPS PUB 46: Data Encryption Standard, 1999.
[3] FIPS PUB 197: Avdanced Encryption Standard, 2001.
[4] FIPS PUB140-2: Security Requirements for Cryptographic Modules, 2001.
[5] L. Afflerbach, Criteria for the assessment of random number generators, Journal of Computational and Applied Mathematics 31 (1990) 3–10.
[6] A. Akhshani, S. Behnia, A. Akhavan, H.A. Hassan, Z. Hassan, A novel scheme for image encryption based on 2d piecewise chaotic maps, Optics Communications 283 (2010) 3259–3266.
[7] R. Anderson, B. Schneier, Description of a new variable-length key, 64-bit block cipher (blowfish), Lecture Notes in Computer Science, vol. 809, Springer, Berlin/Heidelberg, 1994, pp. 191–204.
[8] P. Arora, On the Shannon measure of entropy, Information Sciences 23 (1981) 1–9.
[9] L. Asimow, L. ASA, M. Maxwell, Probability and Statistics with Applications: A Problem Solving Text, Actex Publications, 2010.
[10] A. Atkinson, The computer generation of poisson random variables, Applied Statistics 28 (1979) 29–35.
[11] S. Babbage, C. De Cannière, J. Lano, B. Preneel, J. Vandewalle, Cryptanalysis of sober-t32, in: T. Johansson (Ed.), Fast Software Encryption, Lecture Notes in Computer Science, vol. 2887, Springer, Berlin, Heidelberg, 2003, pp. 111–128.
[12] S. Behnia, A. Akhshani, H. Mahmodi, A. Akhavan, A novel algorithm for image encryption based on mixture of chaotic maps, Chaos, Solitons and Fractals 35 (2008) 408–419.
[13] N. Bourbakis, C. Alexopoulos, Picture data encryption using scan patterns, Pattern Recognition 25 (1992) 567–581.
[14] N. Bourbakis, A. Dollas, Scan-based compression–encryption-hiding for video on demand, IEEE Multimedia 10 (2003) 79–87.
[15] C. Chen, C. Lin, C. Chiang, S. Lin, Personalized information encryption using ecg signals with chaotic functions, Information Sciences 193 (2012) 125–140.
[16] G. Chen, Y. Mao, C.K. Chui, A symmetric image encryption scheme based on 3d chaotic cat maps, Chaos, Solitons and Fractals 21 (2004) 749–761.
[17] X. Chen, S. Kar, D.A. Ralescu, Cross-entropy measure of uncertain variables, Information Sciences 201 (2012) 53–60.
[18] C. Davis, The computer generation of multinomial random variates, Computational Statistics & Data Analysis 16 (1993) 205–217.
[19] D. Davis, R. Ihaka, P. Fenstermacher, Cryptographic randomness from air turbulence in disk drives, in: Proceedings of the 14th Annual International Cryptology Conference on Advances in Cryptology, CRYPTO '94, Springer-Verlag, London, UK, 1994, pp. 114–120.
[20] L. Dorrendorf, Z. Gutterman, B. Pinkas, Cryptanalysis of the random number generator of the windows operating system, ACM Transactions on Information and System Security (TISSEC) 13 (2009) 10.
[21] W. Fang, The characterization of a measure of information discrepancy, Information Sciences 125 (2000) 207–232.
[22] T. Gao, Z. Chen, A new image encryption algorithm based on hyper-chaos, Physics Letters A 372 (2008) 394–400.
[23] R. González, R. Woods, Digital Image Processing, Pearson/Prentice Hall, 2008.
[24] F. James, A review of pseudorandom number generators, Computer Physics Communications 60 (1990) 329–344.
[25] B. Jun, P. Kocher, The Intel Random Number Generator, Cryptography Research Inc., white paper, 1999.
[26] R.C. Kao, L.H. Zetterberg, An identity for the sum of multinomial coefficients, The American Mathematical Monthly 64 (1957) 96–100.
[27] V. Katos, A randomness test for block ciphers, Applied Mathematics and Computation 162 (2005) 29–35.
[28] A. Kumar, M.K. Ghose, Extended substitution–diffusion based image cipher using chaotic standard map, Communications in Nonlinear Science and Numerical Simulation 16 (2011) 372–382.
[29] H.S. Kwok, W.K.S. Tang, A fast image encryption system based on chaotic maps with finite precision representation, Chaos, Solitons and Fractals 32 (2007) 1518–1529.
[30] D. LeBlanc, Statistics: Concepts and Applications for Science, Jones and Bartlett, 2003.
[31] J.L. Leva, A fast normal random number generator, ACM Transactions on Mathematical Software 18 (1992) 449–453.
[32] C. Li, S. Li, G. Alvarez, G. Chen, K. Lo, Cryptanalysis of two chaotic encryption schemes based on circular bit shift and XOR operations, Physics Letters A 369 (2007) 23–30.
[33] S. Li, G. Chen, K. Wong, X. Mou, Y. Cai, Baptista-type chaotic cryptosystems: problems and countermeasures, Physics Letters A 332 (2004) 368–375.
[34] X. Li, A new measure of image scrambling degree based on grey level difference and information entropy, in: Proceedings of the 2008 International Conference on Computational Intelligence and Security, vol. 01, Washington, DC, USA, pp. 350–354.
[35] X. Liao, S. Lai, Q. Zhou, A novel image encryption algorithm based on self-adaptive wave transmission, Signal Processing 90 (2010) 2714–2722.
[36] J. Lizier, M. Prokopenko, A. Zomaya, Local measures of information storage in complex distributed computation, Information Sciences 208 (2012) 39–54.
[37] Y. Mao, G. Chen, S. Lian, A novel fast image encryption scheme based on 3d chaotic baker maps, International Journal of Bifurcation and Chaos in June 14 (2004) 3613–3624.
[38] K. Martin, R. Lukac, K.N. Plataniotis, Efficient encryption of wavelet-based coded color images, Pattern Recognition 38 (2005) 1111–1115.
[39] A. Miranskyy, M. Davison, R. Reesor, S. Murtaza, Using entropy measures for comparison of software traces, Information Sciences 203 (2012) 59–72.
[40] N.K. Pareek, V. Patidar, K.K. Sud, Image encryption using chaotic logistic map, Image and Vision Computing 24 (2006) 926–934.
[41] R. Peck, C. Olsen, J. Devore, Introduction to Statistics and Data Analysis, Duxbury Press, 2004.
[42] J. Peng, D. Zhang, Image encryption and chaotic cellular neural network, Machine Learning in Cyber Trust (2009) 183–213.
[43] C. Petit, F.X. Standaert, O. Pereira, T.G. Malkin, M. Yung, A block cipher based pseudo random number generator secure against side-channel key recovery, in: Proceedings of the 2008 ACM symposium on Information, computer and communications security, ASIACCS '08, ACM, New York, NY, USA, 2008, pp. 56–65.
[44] W. Press, Numerical Recipes: The Art of Scientific Computing, Cambridge University Press, 2007.
[45] J. Rigau, M. Feixas, M. Sbert, Entropy-based adaptive sampling, in: Graphics Interface, pp. 149–157.
[46] B. Schneier, The Twofish Encryption Algorithm: A 128-bit Block Cipher, J. Wiley, 1999.
[47] C.E. Shannon, A mathematical theory of communication, Bell System Technical Journal 27 (1948) 379–423. pp. 623–656.
[48] C.E. Shannon, Communication theory of secrecy systems, Bell System Technical Journal 28 (1949) 656–715.
[49] J. Shen, X. Jin, C. Zhou, A color image encryption algorithm based on magic cube transformation and modular arithmetic operation, Advances in Mulitmedia Information Processing – PCM 2005 (2005) 270–280.
[50] D. Sheskin, Handbook of Parametric and Nonparametric Statistical Procedures, Chapman and Hall/CRC, 2004.
[51] J. Soto, Statistical testing of random number generators, in: Proceedings of the 22nd National Information Systems Security Conference, vol. 10, NIST Gaithersburg, MD, p. 99.
[52] M. Sternstein, Statistics, Barron's Educational Series, 1996.
[53] D. Stinson, Cryptography: Theory and Practice, Chapman and Hall/CRC, 2006.
[54] X. Tong, M. Cui, Image encryption with compound chaotic sequence cipher shifting dynamically, Image and Vision Computing 26 (2008) 843–850.
[55] X. Tong, M. Cui, Image encryption scheme based on 3d baker with dynamical compound chaotic sequence cipher generator, Signal Processing 89 (2009) 480–491.
[56] D. Wagner, B. Schneier, J. Kelsey, Cryptanalysis of the cellular message encryption algorithm, in: B. Kaliski (Ed.), Advances in Cryptology CRYPTO '97, Lecture Notes in Computer Science, vol. 1294, Springer, Berlin/Heidelberg, 1997, pp. 526–537.
[57] Y. Wang, K.W. Wong, X. Liao, G. Chen, A new chaos-based fast image encryption algorithm, Applied Soft Computing 11 (2011) 514–522.
[58] K. Wesolowski, Introduction to Digital Communication Systems, John Wiley & Sons, 2009.

[59] Y. Wu, J.P. Noonan, S. Agaian, Npcr and uaci randomness tests for image encryption, Cyber Journals: Multidisciplinary Journals in Science and Technology, Journal of Selected Areas in Telecommunications (JSAT) (2011) 31–38.

[60] Y. Wu, Y. Zhou, J.P. Noonan, K. Panetta, S. Agaian, Image encryption using the sudoku matrix, in: Mobile Multimedia/Image Processing, Security, and Applications 2010, vol. 7708, SPIE, Orlando, Florida, USA, 2010, pp. 77080P–12.

[61] M. Yang, N. Bourbakis, L. Shujun, Data-image-video encryption, IEEE Potentials 23 (2004) 28–34.

[62] X.Y. Yu, J. Zhang, H.E. Ren, S. Li, X.D. Zhang, A new measurement for image encryption effect, Journal of Physics: Conference Series 48 (2006) 408–411.

[63] L. Zhang, X. Liao, X. Wang, An image encryption approach based on chaotic maps, Chaos, Solitons and Fractals 24 (2005) 759–765.

[64] Q. Zhang, L. Guo, X. Wei, Image encryption using DNA addition combining with chaotic maps, Mathematical and Computer Modelling 52 (2010) 2028–2035.

[65] Y. Zhou, K. Panetta, S. Agaian, C.L.P. Chen, Image encryption using p-fibonacci transform and decomposition, Optics Communications 285 (2012) 594–608.

[66] Z.l. Zhu, W. Zhang, K.w. Wong, H. Yu, A chaos-based symmetric image encryption scheme using a bit-level permutation, Information Sciences 181 (2011) 1171–1186.