Depreciating Motivation and Empirical Security Analysis of Chaos-Based Image and Video Encryption

Mario Preishuber, Thomas Hütter, Stefan Katzenbeisser, *Senior Member, IEEE*, and Andreas Uhl, *Member, IEEE*

***Abstract*— Over the past years, an enormous variety of different chaos-based image and video encryption algorithms have been proposed and published. While any algorithm published undergoes some more or less strict experimental security analysis, many of those schemes are being broken in subsequent publications. In this paper, we show that two main motivations for preferring chaos-based image encryption over classical strong cryptographic encryption, namely computational effort and security benefits, are highly questionable. We demonstrate that several statistical tests, commonly used to assess the security of chaos-based encryption schemes, are insufficient metrics for security analysis. We do this experimentally by constructing obviously insecure encryption schemes and demonstrating that they perform well and/or pass several of these tests. In conclusion, these tests can only give a necessary, but by no means a sufficient condition for security. As a consequence of this paper, several security analyses in related work are questionable; further, methodologies for the security assessment for chaos-based encryption schemes need to be entirely reconsidered.**

***Index Terms*— Chaos-based encryption, image and video encryption, security analysis, cryptanalysis.**

I. Introduction

SINCE Arnold and Avez [1] encrypted the image of a cat by using a chaotic map in 1967, the field of chaos-based image encryption evolved into a lively research area. Inspired by the work of Scharinger and Pichler [2], who applied the Baker map [3] to the discrete case of 2D image encryption, and by the work of Fridrich [4], [5], who extended the discretized map to 3D and composed it with a diffusion mechanism, new chaos-based encryption schemes specifically tailored towards image and video data are being proposed at an almost weekly basis. Consequently, a large number of publications can be observed in recent conferences and journals [6]–[9].

Almost all new chaos-based image and video encryption proposals are motivated by two issues: (i) the potential reduction of computational effort as compared to "naive encryption" (i.e., encryption using a conventional cipher like AES in a proper mode of operation) and (ii) purported security concerns when applying conventional ciphers to images which show significant correlation between adjacent pixels [8], [10], [11]. In this paper we call into question these two main motivations and show that both arguments are actually not correct so that chaos-based image encryption schemes offer little benefits over traditional cryptographic techniques in practice. We acknowledge that special-tailored image and video encryption schemes offer benefits if format-compliance is desired, i.e., the property that even after encryption, a media file still adheres to format requirements and maintains all corresponding functionalities [12]–[14]. However, most chaos-based image encryption schemes are entirely format agnostic, being applied to raw image data and output encrypted raw image data, thereby offering no format compliance. Format compliance is simply not an issue in this field and has never been used to motivate the employment of chaos-based encryption (we acknowledge that a chaos-based scheme used in the context of compression integrated encryption might overall result in a format compliant scheme; still, this is not the reason for its usage in this context).

As for the first claim, the reduction of computational effort, we experimentally show the opposite, i.e., that chaos-based scheme might even be computationally less efficient compared to traditional encryption. We implement some popular chaos-based image encryption schemes and run them on a set of test images. For comparison, we also execute conventional AES encryption as implemented in a cryptographic library on the same machine. In all our experiments, conventional encryption is significantly faster than chaos-based image encryption; due to the availability of highly optimized cryptographic libraries, this situation is unlikely to change in the near future. Our experiments indicate that it is hard for chaotic image encryption implementations to even reach the speed of conventional ciphers.

Regarding the second claim of purported security problems: In many publications it is stated that redundant plaintext data, as naturally found in images, causes security concerns.

While this is true in case no chaining mode of operation (that is, ECB mode) is used, it is common wisdom in cryptography that block ciphers must *always* be used in conjunction with an appropriate mode of operation that chains blocks

Manuscript received August 24, 2017; revised November 30, 2017 and January 20, 2018; accepted January 20, 2018. Date of publication March 5, 2018; date of current version April 26, 2018. This work was supported by the Austrian Science Fund under Grant P27776. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Sherman S.-M. Chow. *(Corresponding author: Andreas Uhl.)*

M. Preishuber, T. Hütter, and A. Uhl are with the Department of Computer Science, University of Salzburg, 5020 Salzburg, Austria (e-mail: uhl@cosy.sbg.ac.at).

S. Katzenbeisser is with the Department of Computer Science, Technische Universität Darmstadt, 64289 Darmstadt, Germany.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIFS.2018.2812080

together or injects additional randomness (such as CBC mode, OFB mode or counter mode). In this case, we expect the encryption to be secure regardless of the data being encrypted; thus, security is also achieved in case sequences with low entropy or high correlation (as seen in real-world images) are encrypted. From a theoretical point of view, typically one demands the security property of ciphertext indistinguishability under a chosen-plaintext attack (IND-CPA) from an encryption system, i.e., a block cipher together with the used mode of operation. Informally, this property requires that, even if an adversary is allowed to query an oracle for encryptions of arbitrary messages, (s)he still cannot distinguish "fresh" encryptions of two messages chosen. Thus, encrypting image or video data with current state-of-the art ciphers (employing an appropriate mode of operation) does not pose a security problem; the argument of low security of conventional ciphers, as put forward by many authors of chaos-based encryption schemes, is flawed.

Therefore, the second motivation to prefer chaos-based encryption for visual data over classical ciphers for security reasons is clearly not a valid one.

Even worse, most published chaos-based image encryption schemes show serious security problems. Most new proposals do not come with a sound security assessment, as common in cryptography. Most authors only attempt to "prove" security experimentally by applying a small set of empirical and/or statistical tests to the encrypted image, e.g., by quantifying correlations or entropy, computing the number of changed pixels, applying sequence tests, or investigating the shape of color value or gray scale histograms. In some papers, resistance against differential attacks, which are typically chosen plaintext attacks, is also studied experimentally, by applying metrics like NPCR and UACI [15]. Furthermore, NPCR is often used to show key sensitivity of an encryption algorithm. To make matters worse, many works use only a limited set of images to derive the results and often only show some graphs to qualitatively "prove" a specific property.

The central problem with assessing security using such metrics is that they are solely computed on the encrypted images as such, and thus do not reflect attackers that utilize knowledge of the encryption algorithm (as demanded by Kerckhoffs' principle) during their attack. Indeed, passing these tests is only a necessary condition for a secure scheme, but not a sufficient one. Moreover, many of these tests do not have an explicit (statistical) decision criterion if an encryption scheme has passed. In fact, in most cases it is only clear what the maximal/optimal value is, but it is not clear under which exact conditions an encryption method passes the test. We only learn if the value is better or worse compared to others (a relative criterion but not an absolute one), and if it is somewhat "close" (without exactly defining the meaning of this term) to the optimum. This fact is also the reason for using the term "empirical security analysis" in the title of the paper but not "statistical".

In this paper we experimentally show that many such security metrics are *insufficient* and can thus *not be used to reason on the security of a cipher at all.* We demonstrate this fact by constructing some trivially breakable encryption schemes and showing that they would be considered "secure" using typical test setups (metrics and parameters) found in papers on chaos-based image encryption. For reference and comparison purposes, we also give test results for some chaos-based image encryption schemes. Indeed, it is worth noticing that many of the proposed schemes have been broken in subsequent publications (e.g. [6], [7], [16]–[21]); it can be expected that these cryptanalysis attempts can be transferred to similar approaches. A central contribution of the paper is thus to experimentally demonstrate that many approaches to "prove" the security of a cipher, which are solely based on statistical properties of the ciphertexts, are fundamentally flawed.

The rest of the paper is structured as follows. Section II gives an overview of some proposed chaos-based encryption schemes as well as cryptanalysis attempts. Section III surveys the encryption schemes we will use for illustrative purposes in the paper, as well as commonly applied security metrics. Section IV presents our experimental results aimed at refuting the two motivations mentioned above. Finally, Section V concludes the paper.

II. Related Work: The Crypto Game in Chaos-Based Image Encryption

In this work we will focus on image encryption techniques, as they also serve as foundations for video encryption schemes operating on a frame-by-frame basis. We distinguish two main classes of encryption techniques that employ two-dimensional chaotic maps: (i) schemes that apply chaotic maps directly to the image itself, represented as matrix, such as [1], [4], and [5], and (ii) schemes that first generate a large pseudo-random stream of bits using a chaos-based random number generator, and subsequently combine image data with the generated stream (for example through the XOR operation), such as [22] and [23]. The latter class is merely a classic stream cipher concept, as it is completely agnostic of the type of data that it encrypts. One option for assessing chaos-based random number generators by statistical means is to apply the NIST test suite [24] proposed for this purpose (see Section III-C). In any case, these chaos-based random number generators are by no means specifically designed for image or video data and should thus not be called image or video encryption schemes. The core focus of this work is thus on class (i).

As mentioned in the Introduction, there are many examples of chaos-based image ciphers that have been proposed in literature, several of them passing experimental security analyses, but nevertheless were broken shortly after. In this section we show some examples for illustrative purposes.

In 2014, Wang and Guo [7] introduced a new encryption algorithm based on chaotic maps. Through experiments they showed that their scheme passes several security tests discussed in the present paper (including entropy, NPCR, histogram analysis and key sensitivity). Due to the results the authors concluded that the algorithm is highly secure. Nevertheless, Yap *et al*. [25] showed in 2015 that the scheme can be broken by a differential attack. The result is remarkable in two ways: First, the attack found by [25] belongs to the class

of known plaintext attacks, and resistance against this class was indeed analyzed experimentally by the original authors. Thus, the experiments provided in the original paper were not able to identify the specific weakness exploited; this calls into question the entire empirical security evaluation methodology. Second, it nicely illustrates that there are weaknesses which manifest themselves only if the attacker exploits knowledge of the encryption algorithm, and cannot easily be found by metrics that operate on encrypted images only.

In 2007, Zang *et al*. [26] introduced a chaos-based image encryption algorithm that shares similarities with a Feistel cipher, where Arnold's cat map is repeatedly used as round function. Again, the authors performed an experimental security analysis by looking at correlations, image histograms and key sensitivity. Even though no attempts were made to cryptanalyze the scheme, i.e., by considering an attacker that exploits knowledge of the structure's cipher, the authors claimed "immunity to many forms of attacks." Zhang *et al*. [6] performed a detailed cryptographic analysis of the scheme. Again, the scheme could be broken by a differential attack; the authors also showed that the key space of the cipher was too small to meet current security standards. Furthermore, the paper also calls into question the second main motivation for chaos-based encryption, namely low encryption times.

In a similar fashion as the two above-mentioned cases, several other multimedia encryption schemes have been broken by advanced cryptanalysis attempts. The design by Yen *et al*. [23], proposed without any security analysis, has been broken by Li *et al*. [20] through a differential attack. The two schemes by Chen *et al*. [27], [28], originally published with only an abstract security argument, were subsequently broken by Li *et al*. [21] by chosen plaintext attacks. The scheme by Feki *et al*. [29], employing a modified Henon map, also published without a thorough security analysis, was subsequently broken by Alvarez *et al*. [16]. A proposal by Yen and Guo [22], which works by using a chaotic system to steer the re-arrangement of bits within every pixel of the image, was completely broken by Li and Zheng [17]. A hierarchical image encryption scheme employing permutation only (HCIE [30]) has been shown to be highly insecure by Li recently [31]. A cipher by Wang *et al*. [8], who applied chaotic sequences to the three color bands of a color image independently and who tested the security of their scheme experimentally using some metrics contained in this paper, was broken in Li *et al*. [32], [33], again by utilizing knowledge of the encryption algorithm in the attack. The proposal by El-Latifa and Niua [11] was broken a year later by Liu and Liu [9] using a known-plaintext attack, requiring only knowledge of one pair of plaintext and ciphertexts. The proposal [34] was broken by [35].

The highly referenced Arnold cat map has been integrated into an image encryption scheme, combined with another chaotic scheme to change the gray values of the shuffled pixels, by Guan *et al*. [36]. This scheme has successfully been cryptanalyzed by Cokal and Solak [37], using known plaintext as well as chosen plaintext attacks, revealing all secret parameters. Finally, also the highly referenced paper by Fridrich [4], [5] was subject to cryptanalysis. The paper proposes an encryption scheme which is based on chaotic confusion and pixel diffusion, performed through several iterations. Analyses of this algorithm have been performed by Lian *et al*. [18] and Solak *et al*. [38]. They conduct a brute force attack, known- and selected plaintext attacks, as well as chosen-ciphertext attacks, demonstrating security problems. The recent paper by Xie *et al*. [39] gives deeper insights into the properties of Solak's attack and provides bases for further optimizing the attacks. General recommendations on the design of chaos-based ciphers were given by [40]; some authors proposed new security metrics as well [41].

In summary, the above-mentioned related work supports the conjectures mentioned in the Introduction: many chaos-based encryption schemes are published either without any security considerations, or they come with an experimental statistical validation of the ciphertexts only. Even in case experiments demonstrate that a cipher passes some security tests, several schemes have been broken through attacks that exploit inherent weaknesses of the encryption algorithms which did not manifest themselves in statistical properties of the ciphertexts. This seriously questions the usefulness of empirical security measures for security assessment. Some authors recently came to a similar conclusion, e.g., Yap *et al*. [25] raised doubts about the appropriateness of NPCR and UACI to assess encryption security, based on their successful cryptanalysis of [7]; and Wu *et al*. [15] base their criticism on statistical considerations concerning the employment of NPCR and UACI. In this paper we will confirm these critics and argue that empirical methods cannot provide a sound security analysis.

III. Encryption Algorithms and Security Assessment Metrics

To foster reproducible research, all software written for this paper, including image encryption techniques, security assessment metrics and the experimental framework, are open source and freely available at GitHub.[1] Software is implemented in C++. We used the CImg library[2] to handle images; furthermore, we use the RC4 algorithm implementation written by Mark Loiseau.[3]

To experimentally test the hypothesis that metrics computed on encrypted images can be used to determine the security of a cipher tailored towards images, we compare common chaos-based encryption schemes to two schemes that are deliberately designed to be insecure. In this section we first review chaos-based image encryption schemes used in this paper, subsequently report on the design of our insecure ciphers and finally review common security metrics.

A. Chaos-Based Image Encryption Schemes

1) Arnold's Cat Map [1]: This chaotic map, shown to be insecure in [37], is an example for a chaotic map, where

[1] https://github.com/mpreis/seth
[2] http://cimg.sourceforge.net
[3] http://markloiseau.com

an image is stretched, cut and reorganized. The generalized transformation Γ is given by

$$\Gamma : \begin{bmatrix} x \\ y \end{bmatrix} \rightarrow \begin{bmatrix} 1 & p \\ q & pq+1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \mod n, \quad (1)$$

where $X = [x\ y]^T$ is a pixel of an $n \times n$ image, with $1 \leq x, y \leq n$, and p, q are positive integer numbers used as key. Furthermore, the algorithm can be repeated t times, t being part of the key as well.

For encryption with Arnold's cat map we apply a randomly chosen number of iterations. According to [42] we have chosen an upper bound of 45 iterations; in order to avoid visible structures caused by low iteration count (see Fig. 1(b)), we set the lower bound to 10. The parameters p and q are also randomly selected from the interval $[10, 45]$.

2) Baker's Map [3]: This chaotic map, shown to be insecure in [18] and [38], is probably the best known chaotic map. An image is split vertically, stretched horizontally and then the resulting pieces are stapled on top of each other. The number of splits, as well as the position of the splits, can be chosen arbitrarily, determined by a key.

This map can be applied to an image as follows [4], [5]: Define a sequence $n_1, n_2, \ldots, n_k$, where k is the number of rectangles the image is split into. Each n_i must divide the image width N without remainder and $n_1 + \cdots + n_k = N$. Furthermore, let $N_i = n_1 + \cdots + n_i$ and $N_0 = 0$.

Consider a pixel (r, s) with $N_{i-1} \leq r < N_i$ and $0 \leq s < N$ in an $N \times N$ image. This pixel (r, s) is mapped to

$$B(r,s) = \Big(\big(q_i \cdot (r - N_i) + (s \bmod q_i) \big) = \Big(\frac{s - (s \bmod q_i)}{q_i + N_i} \Big) \Big), \quad (2)$$

where $q_i = N/n_i$. So far, the algorithm is just a permutation of pixels. To distribute the gray values, a substitution step can be added. In particular, the pixel (r, s) with gray value g_{rs} gets mapped to a pixel at position $B(r, s)$, and its gray value is changed to $h(r, s, g_{rs})$, dependent on the pixels new position and the old gray value. A popular choice is the function

$$h(r, s, g_{rs}) = (g_{rs} + r \cdot s) \mod L, \quad (3)$$

where L is the number of gray values.

Like Arnold's cat map, Baker's map may be applied several times as well. We choose the number of iterations randomly between 10 and 45. To determine the number of slices we generate a set of n random numbers until the sum of these numbers is equal to or greater than the width of the image. If the sum is greater than the image width, the last value is replaced by the image width minus the sum of the $n - 1$ previous values. Each number indicates the width of a single slice.

Fig. 1 illustrates the outcome of applying all variants of chaos-based ciphers described above to the Lena image, namely Arnold's cat map, Baker's map and Baker's map with substitution. The stripe pattern in the image encrypted using Arnold's cat map (Fig. 1(b)) originates from the low number (4) of iterations applied (not admissible in our experimental parameter setting).

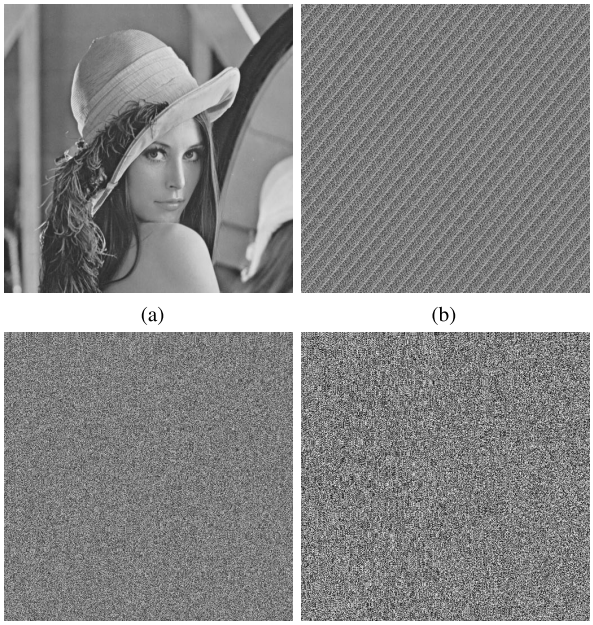

Fig. 1. Lena image encrypted with chaos-based encryption variants. (a) Original. (b) Arnold's cat map. (c) Bakers's map. (d) Bakers's map w. substitution.

3) Two-Dimensional Logistic Chaotic Map [43]: The logistic map is another popular chaotic map that has been used to design image encryption techniques. A direct and exclusive application of the logistic map in image encryption [34], [44] has been shown to be insecure [45]. One of the more recent approaches [43], which we employ in our experiments, is based on a 2D logistic map combined with other encryption stages. So far, no successful cryptanalysis has been published. The algorithm has been used as a reference to be compared to, e.g. in [46] it shows decent behaviour in statistical tests also conducted in this paper and in [47] it is used in a comparison wrt. noise and data loss attacks. The key consists of 256 bits, which are split into four 52 bit and eight 48 bit values, describing the initial value and the parameter of the map. The encryption happens in three phases:

- *2-D Permutation:* Rows and columns are permuted by applying 1-D logistic permutation to rows and columns consecutively.
- *2-D Diffusion:* For each 4x4 block of the plaintext image a multiplication with maximum distance separation matrices computed from random permutation matrices over a finite field $GF(2^8)$ is computed. We have used the finite field implementation of crypto++.[4]
- *2-D Transposition:* Subfunctions are applied to every element of each 4x4 block of a preprocessed version of the input to this stage and later added to the plaintext image.

For detailed information we refer to [43]; a MATLAB implementation of the algorithm by the authors is available as well.[5]

[4] http://www.cryptopp.com

[5] https://sites.google.com/site/tuftsyuewu/source-code

B. Deliberately Insecure Encryption Schemes

As a deliberately insecure cipher we use a stream cipher, which computes the XOR of the output of a source of pseudo-randomness and the image content. We distinguish two modes, named pixel-mode and MSB-mode:

- In pixel-mode, an image is encrypted pixel by pixel, starting at the most significant bit up to the least significant bit.
- In MSB-mode, all the most significant bits of the image are encrypted first, then the second most significant bits, and so on.

We use two different ways to generate the pseudo-random stream: One uses the insecure cipher RC4 and one uses a linear congruential random number generator, which is predictable. Fig. 2 illustrates the four encryption schemes when applied to the Lena image.

1) XOR OTP RC4: The XOR OTP RC4 approach uses a string as key for the RC4 algorithm. This key has a maximum size of 256 characters, where each character is a randomly chosen ASCII character. The output of the RC4 cipher are values between 0 and 255, which are used as pseudo-random stream. The XOR OTP has a very large key space of size 256^{256}, but this does not allow us to conclude that there are no security issues: The vulnerability stems from the RC4 algorithm itself, which is known to be insecure [48].

2) XOR OTP CSTD: The second approach uses the pseudo random number generator of the C Standard Library and is called XOR OTP CSTD. The generator is based on a simple (single state) linear congruential generator (LCG). The key is an integer of 32 bits, which is used as seed. This cipher has a very small key space of 2^{32} elements, which can easily be brute-forced. Thus, this scheme has to be considered highly insecure as well.

C. Security Assessment Metrics

In this section we describe well known security assessment metrics that are used in the majority of papers on chaos-based image and video encryption to experimentally demonstrate the security of new ciphers. We have chosen the tests by analyzing the experimental section of several papers [34], [36], [49]–[64].

1) Correlation: A popular measure is the correlation between pixels in horizontal, vertical and diagonal direction. For this purpose, one chooses N pairs of pixels (x, y) which are adjacent in horizontal, vertical and diagonal direction, and computes the correlation coefficient between the gray values

$$r_{xy} = \frac{\frac{1}{N}\sum_{i=1}^{N}(x_i - E(x)) \times (y_i - E(y))}{\sqrt{\frac{1}{N}\sum_{i=i}^{N}(x_i - E(x))^2}\sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - E(x))^2}}, \quad (4)$$

where $E(x) = \frac{1}{N}\sum_{i}^{N} x_i$. The correlation coefficient r_{xy} is a value between -1 and 1, where 1 and -1 indicates a high correlation and 0 no correlation. Because neighboring pixels in images are highly correlated, we expect high values when applying the metric to original images; however, to avoid statistical attacks, correlation values should be around zero for encrypted images. There is no clear (statistical) decision criterion for passing this test.

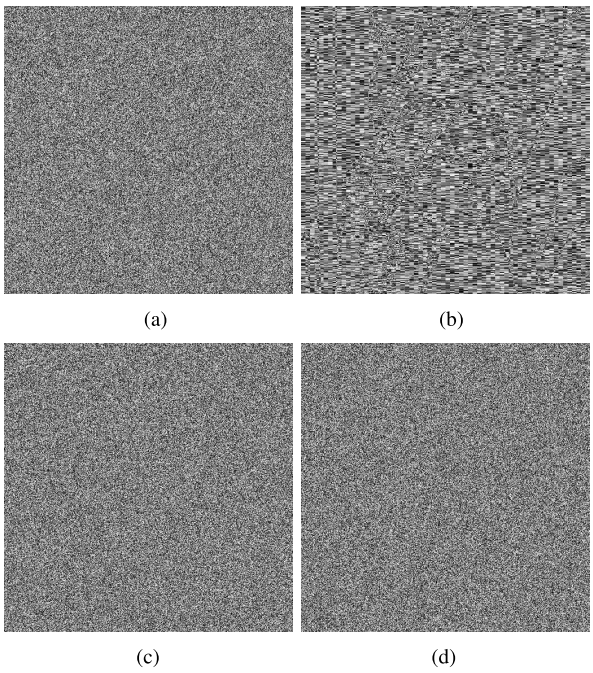

Fig. 2. Lena image encrypted with insecure ciphers. (a) XOR OTP RC4 pixel. (b) XOR OTP RC4 MSB. (c) XOR OTP CSTD pixel. (d) XOR OTP CSTD MSB.

2) Entropy: The Shannon entropy

$$H(m) = -\sum_{i=0}^{N} p(m_i) \times \log_2 p(m_i), \quad (5)$$

where $p(m_i)$ is the probability that a pixel has a specific gray value m_i and N denotes the total number of gray values, is also often used to measure the "randomness" of an encrypted image. In practice, $p(m_i)$ is estimated by the fraction of pixels with gray value m_i to the total number of image pixels. A completely random 8-bit gray scale image would achieve an entropy of 8 bits; an encrypted image should thus get close to that value. There is no clear (statistical) decision criterion for passing this test.

3) Gray Scale Histogram Variance: The (color or gray scale) histogram of an image is typically highly non-uniform. In contrast, a properly encrypted image should have a histogram that is close to uniform. Some papers use the variance of the number of entries of the gray value histogram bins to measure how uniform the histogram of an image is. A value of 0 corresponds to a totally uniformly distributed histogram, which would be the optimum for an encrypted image. There is no clear (statistical) decision criterion for passing this test.

4) Number of Pixel Change Rate (NPCR): The relative number of different pixels between two images I_1 and I_2 of size $N \times M$ is calculated by

$$\text{NPCR} = \frac{\sum_{i=1}^{N}\sum_{j=1}^{M}\delta_{I_1(i,j),I_2(i,j)}}{NM} \times 100\%, \quad (6)$$

where $\delta_{x,y}$ is the Kronecker delta, i.e., $\delta_{x,y} = 1$ if $x = y$ and $\delta_{x,y} = 0$ otherwise. NPCR is used to see how much the original image I_1 differs from the encrypted version I_2;

in addition, it is sometimes used for the key sensitivity test explained below. The higher the value, the better for security, with a maximum of 100%. There is no clear (statistical) decision criterion for passing this test.

5) Unified Average Changing Intensity (UACI): Like NPCR, UACI is also used to measure the difference between two images I_1 and I_2 of size $N \times M$ and is calculated by

$$\text{UACI} = \frac{1}{NM} \sum_{i=1}^{N} \sum_{j=1}^{M} \frac{I_1(i,j) - I_2(i,j)}{\text{tonal range}} \times 100\%, \quad (7)$$

where the tonal range is 255 in our case.

UACI is also used to see how much an encrypted image I_2 differs from its original image I_1. Higher numbers indicate better results, while the ideal value highly depends on the tonal range and the grey value distribution in the considered image. Note that in case of applying UACI in the context of key sensitivity tests comparing two encrypted images, the optimum value is 33.3%. However, we only apply UACI in the traditional way comparing original and encrypted image where the value is found to be typically lower. There is no clear (statistical) decision criterion for passing this test.

6) Key Sensitivity, Differential Attacks: Any ideal cipher should have the property that decrypting a ciphertext using a slightly different key than the one used for encryption should lead to a completely different result. A common way in the literature on chaos-based image encryption to validate this property is an application of NPCR: First, an image is a encrypted with key k_1. The encrypted image is subsequently decrypted with a slightly different key k_2. The difference between k_1 and k_2 may, for example, be one bit flip at a random position (this is exactly the way how we employ this test in the experiments). Now, the NPCR between the original image and the decrypted image is computed. Ideally, NPCR should reach 100%.

Some authors performed similar steps to measure strength of the cipher against differential attacks: One encrypts the same image under two different keys and measures the similarity of both ciphertexts using NPCR. Wu *et al.* [15] provide a statistical decision criterion to judge if an algorithm passes this test.

7) Sequence Tests: Another common method to evaluate image encryption schemes is to test whether encrypted images are "random" by common statistical tests, such as the sequence test. The single bit test checks if the number of zeros n_0 of an image is equal to the number of ones n_1. In a random string (and a well-encrypted image) these two numbers should be roughly equal; this condition can be checked using a chi-square test statistic

$$\chi^2 = \frac{(n_0 - p_0 * n)^2}{p_0 * n} + \frac{(n_1 - p_1 * n)^2}{p_1 * n}, \quad (8)$$

where $p_0 = p_1 = 1/2$ are the ideal probabilities for observing zeros and ones and $n = n_0 + n_1$ denotes the total number of bits in the encrypted image. The test can be extended in a straightforward way to check the distribution of sequences of bits 00, 01, 10 and 11 in the image using the statistic

$$\chi^2 = \frac{(n_{00} - p_{00} * n)^2}{p_{00} * n} + \frac{(n_{01} - p_{01} * n)^2}{p_{01} * n} + \frac{(n_{10} - p_{10} * n)^2}{p_{10} * n} + \frac{(n_{11} - p_{11} * n)^2}{p_{11} * n}, \quad (9)$$

where again $p_{00} = p_{01} = p_{10} = p_{11} = 1/4$ denote the expected probabilities for observing the four sequences and $n_{00} + n_{01} + n_{10} + n_{11} = n/2$ denote the number of observed bit tuples, which sum up to the total number $n/2$ of tuples.

Depending on the significance value α, an image will pass or fail the single or double sequence test. To pass, the calculated value χ^2 must be smaller than the given value from the chi-square test.

8) NIST (Pseudo) Random Number Generators Test Suite [24]: NIST provides software and a corresponding documentation for validation of (pseudo) random number generators for cryptographic applications using a statistical test suite, implementing a wide range of 15 different statistical tests measuring various distribution properties of binary sequences. The documentation states: "*These tests may be useful as a first step in determining whether or not a generator is suitable for a particular cryptographic application. However, no set of statistical tests can absolutely certify a generator as appropriate for usage in a particular application, i.e., statistical testing cannot serve as a substitute for cryptanalysis.*" Thus, it gets clear that the authors do consider passing these test as a prerequisite for a sensible (pseudo) random number generator, but not as a sufficient criterion for its security. Some (not many) image encryption proposals have been evaluated using this NIST test suite – in some rare cases [65]–[68] by directly applying it to encrypted images (which is at the core of this paper), more often in order to assess chaotic binary sequences which are subsequently employed in image encryption techniques (e.g. [69], which is the application setting this suite has been designed for). Following the intentions of e.g. [65]–[68], we have applied the NIST test suite to our data set as follows: Each image from our dataset is encrypted using 10 randomly chosen keys; the resulting data is then subjected to the test suite analysis. Each encrypted image is transfered into a NIST *bitsequence* by reading the pictorial information from each pixel, bit after bit.

IV. Experimental Evaluation

The goal of this section is twofold. First, we demonstrate that encryption algorithms deliberately chosen to be insecure, as the ones described in Section III-B, perform well and/or pass the set of security tests of Section III-C, which are commonly used to validate the strength of chaos-based encryption algorithms. This demonstrates that these metrics *cannot* be used to properly evaluate the security of image encryption schemes. Furthermore, we also experimentally demonstrate that chaos-based image encryption schemes do not outperform classic encryption algorithms in terms of computational efficiency (in contrast, our implementations turn out to be significantly slower as compared to several variants of AES-based encryption), which invalidates the second key argument for the development of specially tailored image ciphers.

In all our experiments we used the images of the USC-SIPI image database,[6] maintained by the University of Southern California, and a dataset of standard test images maintained by the University of Granada,[7] overall 128 images. The used images are of size 512×512 and 8bpp gray scale.

In the following section we use shortcuts for the encryption algorithms. We denote the XOR OTP RC4 scheme in pixel-mode as *xor-otp-pix* and the MSB-mode as *xor-otp-msb*. If the algorithm uses the C Standard Library random number generator, the used terms are *xor-otp-cstd-pix* and *xor-otp-cstd-msb*. Arnold's cat map algorithm is simply named *arnold* and Baker's map is named *baker*. The substitution-mode of the Baker's map is called *baker-sub* while the 2-D logistic map encryption is called *2d-log-map(-256)*.

A. Security Evaluation Results

In the following paragraphs we provide quantitative results for all the metrics described in Section III-C. All test images were encrypted 100 times using different randomly chosen keys. Results, averaged over all images and all 100 trials, are summarized in Table I and discussed below. Where appropriate and interesting we also provide qualitative results (encryptions of the Lena image). It is important to note that the results of statistical tests highly depend on the variant of the test used and the parameters chosen in the tests. When changing parameters, one might get completely different results. The test parameters in this work are chosen in a way to ensure the best possible comparability to test results given in the literature. Thus, we select parameters (i) according to mandatory requirements (like image size and bit depth) and (ii) to match settings that have been used in the majority of previous employments of a certain test. For each test, we will explicitly state the parameters used.

1) Correlation: We use all available overlapping pixel pairs in each image using a zig-zag scan, i.e., $N = 262143$. We start by considering qualitative results. For simplicity we only display results for vertically adjacent pixels; other dimensions (horizontal and diagonal) show very similar results. Fig. 3 shows correlations present in different encryptions of the Lena image. As expected, there is a strong correlation of pixels in the original image, which is indicated by the clustering of the plotted points along a diagonal. Arnold's cat map shows much better correlation properties when compared to the original image; nevertheless we still see areas with a much higher concentration of points than others. Furthermore, not the whole spectrum of possible correlation values is exploited. Using Baker's map for encryption does not lead to better results. These observations alone show that neither Arnold's nor Baker's map properly encrypt images. The situation only changes if one employs Baker's map in substitution-mode: The points in the scatter plot are distributed over the whole value spectrum and seem to be uniformly distributed; in addition, there are no clusters of points (this is also the case for the 2-D logistic map). Using the XOR OTP CSTD and XOR OTP RC4 algorithms for encryption results in very low correlation

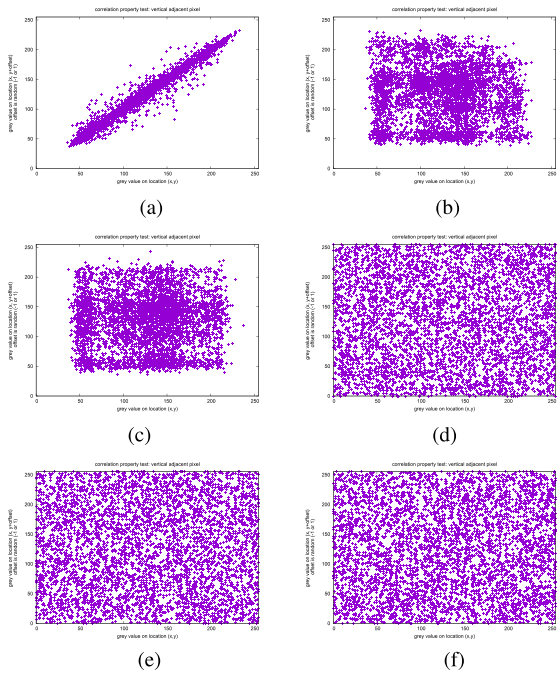

Fig. 3. Vertical correlation of the Lena image: (a) original image, and image encrypted (b) by Arnold's cat map, (c) by Baker's map, (d) by Baker's map in substitution mode, (e) by the XOR OTP CSTD algorithm in pixel-mode, (f) by the XOR OTP RC4 algorithm in pixel-mode.

in encrypted images, despite the ciphers being highly insecure. A visual inspection would thus mistake two insecure ciphers as purportedly secure.

These preliminary observations are confirmed by the results in Table I. Values for vertical correlation of Arnold's cat map and Baker's map are slightly higher than for the two insecure ciphers. Nevertheless, all show rather low correlation except for the XOR OTP RC4 in MSB mode, in many cases similar or even better compared to the so far unbroken 2-D logistic map encryption scheme. Thus, considering this measure, one would potentially mistake at least three out of four insecure ciphers as "secure".

2) Entropy: For our 8bpp images, $N = 255$. Entropy values of both insecure ciphers as well as Baker's map in substitution mode and the 2-D logistic map are almost identical and close to the optimal value of 8. The values for Arnold's cat map and Baker' map tend to be clearly smaller and also exhibit rather high standard deviation. Again, our insecure ciphers achieve near optimal values, clearly indicating to pass this test.

3) Gray Scale Histogram: For computing the histogram variance, we use histograms with 256 bins according to the number of gray scales in our images. We compute the mean of the number of bin entries and subsequently employ the variance of this expression as our histogram "metric". We first perform a qualitative visual inspection. Figure 4 depicts the histogram of the original Lena image and histograms of encryptions using different image encryption schemes. As expected, the gray scale values are not uniformly distributed in the original image. The histograms of the Arnold's cat map and Baker's map encrypted images look exactly like the histogram of the original image. The reason for this behavior

[6]http://sipi.usc.edu/database/

[7]http://decsai.ugr.es/cvg/CG/base.htm

TABLE I

SECURITY METRICS FOR DIFFERENT ENCRYPTION SCHEMES. METRICS ARE AVERAGED OVER ALL IMAGES AND TRIALS WITH 100 DIFFERENT RANDOMLY CHOSEN KEYS, MEAN AND STANDARD DEVIATION (THE LATTER IN BRACKETS) ARE GIVEN

Encryption	Correlation Vertical	Correlation Horizontal	Correlation Diagonal	Entropy	Histogram bin variance	UACI	NPCR	Key sensitivity NPCR
xor-otp-pix	-0.000392	-0.000635	-0.000599	7.999298	3.821070	31.947407	99.609157	99.608766
	(0.014474)	(0.014396)	(0.014340)	(0.000060)	(5.701182)	(3.768387)	(0.012013)	(0.012860)
xor-otp-msb	-0.000041	0.691057	0.000033	7.998582	7.905122	31.949321	99.609191	99.601594
	(0.014603)	(0.142024)	(0.014821)	(0.000919)	(13.847178)	(3.770357)	(0.034536)	(0.050401)
xor-otp-cstd-pix	0.000240	0.000706	-0.000198	7.999298	3.938914	31.950682	99.608912	99.609563
	(0.014264)	(0.014483)	(0.013817)	(0.000066)	(5.326925)	(3.769676)	(0.012356)	(0.012007)
xor-otp-cstd-msb	0.000073	-0.000295	-0.000231	7.999298	4.069816	31.948473	99.609719	99.609563
	(0.014266)	(0.014067)	(0.014319)	(0.000063)	(6.044997)	(3.768803)	(0.011793)	(0.012204)
baker-sub	0.002311	-0.000624	0.000868	7.998946	6.354168	31.949338	99.608359	99.515037
	(0.031030)	(0.014338)	(0.014396)	(0.000571)	(9.909429)	(3.767817)	(0.012387)	(0.638901)
baker	0.005000	-0.000193	0.000018	6.272438	2076414.982431	19.517311	97.056005	96.967023
	(0.039854)	(0.014994)	(0.014818)	(1.043449)	(18635101.944347)	(8.427148)	(7.290827)	(7.326533)
arnold	-0.007277	-0.003278	-0.001093	6.279706	2108514.488736	19.421496	96.642943	96.642943
	(0.125877)	(0.131424)	(0.096666)	(1.046640)	(18907240.823312)	(8.438301)	(8.390801)	(8.390801)
2d-log-map	0.000188	-0.000022	0.000239	7.999300	4.158525	31.952216	99.609335	90.114524
	(0.014048)	(0.014365)	(0.014646)	(0.000063)	(6.295641)	(3.768127)	(0.012245)	(28.737761)
ideal value	0.0	0.0	0.0	8.0	0.0	n.a.	100.0	100.0

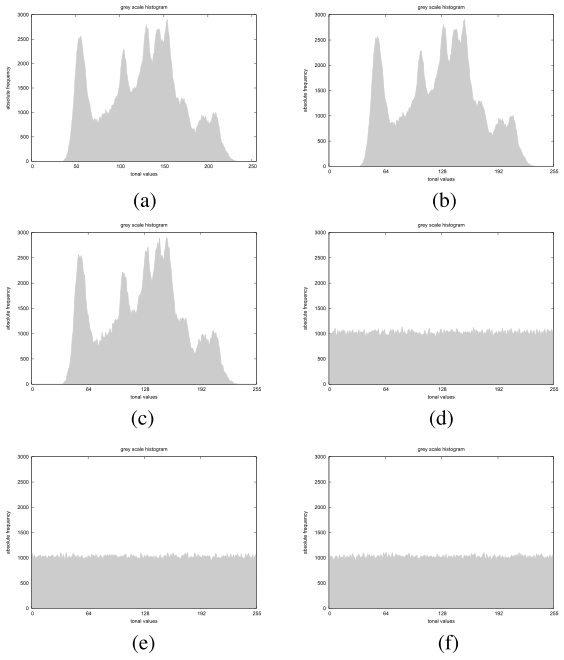

Fig. 4. Gray scale histogram of the Lena image: (a) original image, and image encrypted (b) by Arnold's cat map, (c) by Baker's map, (d) by Baker's map in substitution mode, (e) by the XOR OTP CSTD algorithm in pixel-mode, (f) by the XOR OTP RC4 algorithm in pixel-mode.

is that both algorithms do not change the gray scale values, but only apply a permutation on the original image. This is a severe security issue, which can potentially be mitigated with substitution mode; as expected, the image histogram of an encryption using Baker's map in substitution mode is much more uniform (the same is true for the 2-D logistic map, not shown). However, this is again no clear indicator for security, as our two insecure ciphers yield to an almost perfectly flat gray scale histogram as well.

This qualitative analysis corresponds to the numbers given in Table I. The variance of the histogram bins is very small (top performing and close to the optimum, indicating a near uniform distribution) when using both XOR OTP PIX variants and XOR OTP CSTD in MSB mode. The 2-D logistic map follows closely, and Baker's map in substitution mode is clearly worse. Again, XOR OTP RC4 in MSB mode is the worst of these five ciphers (compare also poor horizontal correlation values discussed before). Note that these observations wrt. XOR OTP RC4 in MSB mode correspond to the visualization in Fig. 2(b), where more structure is observed compared to the other "insecure" ciphers. Mean values and standard deviation of the histogram bin variance for Arnold's cat map and Baker's map is extremely high as to be expected from Fig. 4(b)(c). Again, a small histogram bin variance is thus no clear indicator for security as it is met at least with three out of four insecure ciphers.

4) NPCR and UACI: In our tests, according to the image size, $N = M = 512$ and *tonal range* is set to 255. Table I shows the quantitative results of the NPCR test. The values for our four insecure ciphers are above 99%, which is comparable to the best NPCR values found in the literature on chaos-based ciphers. In particular, this value is virtually identical to that of the so far unbroken 2-D logistic map encryption scheme. The NPCR value for Arnold's cat map and Baker's map is smaller; further, the standard deviation of Arnold's map and Baker's map is much larger. Similar results are obtained for the UACI test as well. Baker's map and Arnold's cat map both stay under 20 percent, these are the worst results. Again, XOR OTP RC4 and XOR OTP CSTD perform very well at the UACI test (again virtually identical values compared to the 2-D logistic approach) clearly demonstrating that excellent performance in NPCR or UACI is not an indicator for a secure cipher.

5) Key Sensitivity Test: As stated above, we decrypt an encrypted image with an invalid key, differing only in one bit, and compute the NPCR measure between the original and the decrypted image. With a proper encryption scheme one should not be able to recognize the original image; thus, we expect a NPCR of close to 100%. Thus, I_1 is the original image, I_t is the image resulting from encrypting I_1 with a random key. Then we flip one bit of this key at a random position, and

TABLE II
SECURITY METRICS FOR DIFFERENT ENCRYPTION SCHEMES, CONTINUED

Encryption	χ^2 sequence test		Relative number of bits		Relative number of tuples			
	single bit	double bit	0	1	00	01	10	11
xor-otp-pix	0.968275	3.032797	0.499982	0.500018	0.249998	0.249991	0.249975	0.250035
	(1.354410)	(2.356046)	(0.000339)	(0.000339)	(0.000422)	(0.000415)	(0.000439)	(0.000422)
xor-otp-msb	3.076486	9.442041	0.500033	0.499967	0.250023	0.249999	0.250021	0.249957
	(4.898279)	(8.700496)	(0.000605)	(0.000605)	(0.000753)	(0.000766)	(0.000731)	(0.000749)
xor-otp-cstd-pix	0.984671	3.012089	0.499990	0.500010	0.249991	0.249996	0.250001	0.250012
	(1.366893)	(2.500854)	(0.000343)	(0.000343)	(0.000417)	(0.000434)	(0.000423)	(0.000421)
xor-otp-cstd-msb	0.945347	2.972037	0.499998	0.500002	0.249998	0.250000	0.250001	0.250002
	(1.367855)	(2.347519)	(0.000336)	(0.000336)	(0.000416)	(0.000421)	(0.000434)	(0.000412)
baker-sub	1.647303	5.058866	0.499991	0.500009	0.249990	0.250006	0.249996	0.250008
	(2.339221)	(4.864916)	(0.000443)	(0.000443)	(0.000542)	(0.000555)	(0.000550)	(0.000551)
baker	40566.758080	94071.390496	0.521644	0.478356	0.267148	0.246160	0.262832	0.223859
	(119410.852709)	(215957.063751)	(0.066113)	(0.066113)	(0.079442)	(0.063805)	(0.065881)	(0.081030)
arnold	40894.031006	94106.872902	0.522039	0.477961	0.267881	0.245816	0.262500	0.223803
	(122175.966647)	(218881.476244)	(0.066277)	(0.066277)	(0.079151)	(0.063708)	(0.065654)	(0.081484)
2d-log-map	1.049011	3.078875	0.500002	0.499998	0.249999	0.250004	0.250003	0.249995
	(1.467049)	(2.480322)	(0.000354)	(0.000354)	(0.000418)	(0.000422)	(0.000430)	(0.000444)
ideal value	≤ 3.84	≤ 7.81	0.5	0.5	0.25	0.25	0.25	0.25

decrypt I_t into I_2 with this key. Corresponding to the image size, $N = M = 512$.

The 2-D logistic map approach exhibits the worst average result in this test, also with an extremely high standard deviation, which indicates a problem in key scheduling. Arnold's cat map and Baker's map do lead to a result close to 97%, but compared with the other algorithms these results are low (in addition they show a high standard deviation). Baker's map in substitution-mode shows an improvement, but still worse compared to the insecure ciphers in terms of mean and standard deviation. Again, our insecure XOR OTP RC4 and XOR OTP CSTD algorithms lead to the best results. Both algorithms (in MSB-mode and pixel-mode) result in an NPCR value over 99.6 percent. Following the classification of Wu *et al.* [15], who consider three levels of significance for this test in their work, all XOR OTP variants pass the test at all three levels of significance considered, while the chaotic encryption variants all fail.

6) Sequence Test: In our experiments we chose a significance value of $\alpha = 0.05$ for the chi-square statistic and, according to image size (512^2 pixels) and bit depth (8bpp), $n = 2097152$. The number of degrees of freedom is one for the single bit and three for the double bit test. This means, to pass the single bit test, the χ^2 value has to be lower than 3.84, and to pass the double bit test the limit is 7.81.

Table II shows that most of the encryption schemes pass this test: Almost all XOR OTP variants, Baker's map in substitution mode as well as the 2-D logistic map technique. In addition, the same encryption techniques failing at the single bit test also fail at the double bit test. All others pass both tests. There is one exception: XOR OTP RC4 in MSB-mode does only pass the single bit test (and not with flying colors, though), an observation confirming already seen weaknesses (e.g. wrt. horizontal correlation and histogram bin variance). As expected, Arnold's cat and Baker's map fail these tests in spectacular manner, because they do not change any gray scale values. The XOR OTP RC4 algorithm in pixel-mode achieves the best result. The table also shows the distributions of individual bits and tuples of bits; all deliberately insecure schemes as well as the Baker's map in substitution mode and the 2-D logistic map pass this test while Arnold's cat map and Baker's map fail with clearly worse means and standard deviations.

7) NIST (Pseudo) Random Number Generators Test Suite [24]: Table III reports the ratio of passed NIST tests (averaged over all images and used keys for each encryption technique), as read out from the corresponding *stats.txt* files. The software uses the tests' p-values to determine if a test is passed; we do not change the pre-set default p-values for this assessment and determine the ratio of passed tests in this manner. Furthermore, all other parameters in the software are left at their default values in order to facilitate best-possible comparability to results published earlier, which are also based on these default settings.

Overall, we note that Arnold's cat map and Baker's map exhibit very low passing ratios (and even some 0.0 values for *Approximate Entropy and Universal* and 10 out of 15 tests exhibit passing ratios < 0.1). For the XOR OPT variants we observe much better results for the PIX variants (this is explained by the match between encryption order and chosen NIST data input strategy directly exhibiting the weaknesses of the underlying stream ciphers when considering the MSB modes). The two (or one of the two) XOR OPT PIX algorithms often exhibit the highest overall test passing ratio(s), i.e., for 10 out of 15 tests the deliberately insecure schemes provide the highest (and indeed very high in absolute terms) test passing ratios. For all but *FFT and Serial* tests both XOR OPT PIX algorithms attain passing ratios well above 0.98 which are considered to be excellent values [24]. Even one of the XOR OPT MSB variants is able to achieve excellent passing ratios > 0.98 for some tests, e.g., the *Cumulative Sums, Random Excursions Variant, Linear Complexity and Frequency* tests. The 2-D logistic map technique achieves the best overall result with passing ratios > 0.98 for all tests considered while Baker's map in substitution mode exhibits significant weaknesses in 3 out of 15 tests.

TABLE III

RATIO OF PASSED NIST SECURITY TESTS FOR DIFFERENT ENCRYPTION SCHEMES

	Random Excursions	Cumulative Sums	Random Excursions Variant	FFT	Overlapping Template	Runs	Rank	Longest Run
2d-log-map-256	0.989201	0.990625	0.991301	0.994531	0.980469	0.989062	0.988281	0.985938
baker	0.694444	0.029297	0.979938	0.032031	0.013281	0.086207	0.495312	0.025781
arnold	0.567857	0.023438	0.980952	0.007031	0.014063	0.333333	0.490625	0.025781
baker-sub	0.987561	0.926953	0.990426	0.126562	0.942969	0.963079	0.972656	0.938281
xor-otp-rc4-msb	0.926417	0.771484	0.984173	0.000000	0.010937	0.879742	0.138281	0.000000
xor-otp-cstd-msb	0.439041	0.983594	0.952721	0.100000	0.000000	0.042188	0.167187	0.000000
xor-otp-rc4-pix	0.990048	0.992578	0.992451	0.989844	0.989062	0.991406	0.987500	0.985156
xor-otp-cstd-pix	0.989196	0.991406	0.991126	0.259375	0.981250	0.992188	0.990625	0.984375

	Block Frequency	Approximate Entropy	Non Overlapping Template	Linear Complexity	Serial	Frequency	Universal
2d-log-map-256	0.987500	0.990625	0.990097	0.993750	0.990625	0.991406	0.986719
baker	0.339062	0.000000	0.078326	0.979688	0.000000	0.029687	0.000000
arnold	0.342187	0.000000	0.100560	0.977344	0.007812	0.023438	0.000000
baker-sub	0.977344	0.025781	0.916258	0.987500	0.000000	0.939063	0.950000
xor-otp-rc4-msb	0.000000	0.000000	0.001156	0.990625	0.000000	0.851562	0.000000
xor-otp-cstd-msb	0.796875	0.000000	0.030395	0.840625	0.000000	0.992969	0.000000
xor-otp-rc4-pix	0.988281	0.988281	0.990192	0.985938	0.990234	0.992969	0.991406
xor-otp-cstd-pix	0.990625	0.951562	0.990266	0.989062	0.911328	0.994531	0.989062

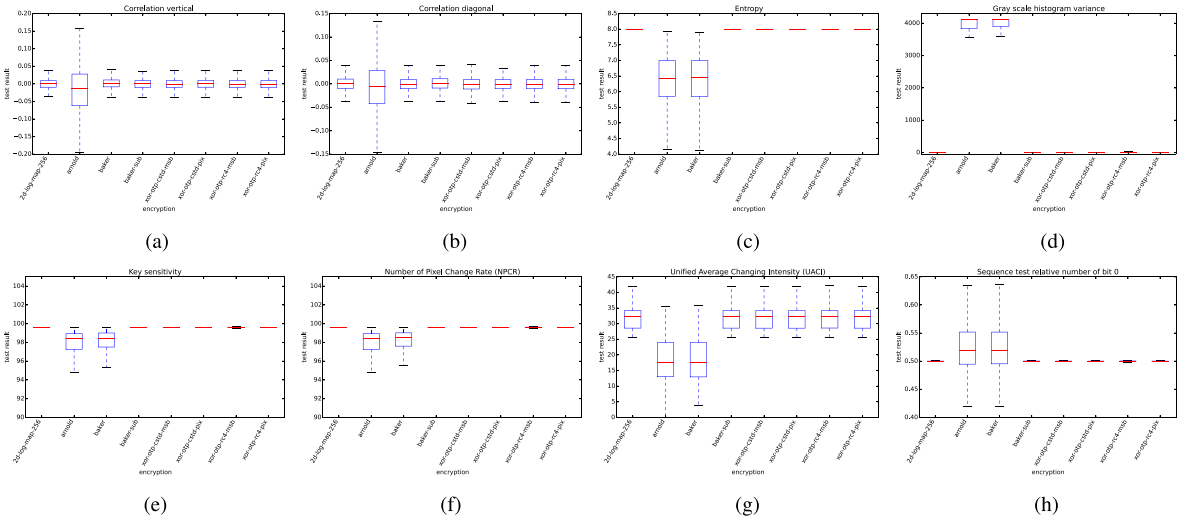

Fig. 5. Boxplots giving statistical information for some of the metrics considered in this paper, computed over the entire data set of test images and keys as specified. (a) Vertical Correlation. (b) Diagonal Correlation. (c) Entropy. (d) Histogram variance. (e) Key sensitivity. (f) NPCR. (g) UACI. (h) Relative number of 0 s.

Thus, achieving high ratios in passing the NIST test suite is not at all a proof for the security of an image encryption scheme as clearly demonstrated by the excellent values achieved in many tests by the XOR OPT PIX algorithms (and in some tests by the XOR OPT MSB algorithms).

8) Conclusions on Security Tests: In summary, our insecure ciphers either passed and/or performed very well in almost all tests – in most cases even with better results than all considered chaos-based image encryption schemes including the so far unbroken 2-D logistic map approach. This can also be seen in Figure 5, which presents statistical data (median as red line and 25% as well as 75% quantiles as boxes content) for some selected metrics considered in this paper, computed over the entire data set of test images. Our results clearly show that none of the above metrics can be used to test security of a cipher for images since the ciphers deliberately chosen to be insecure either pass and/or perform very well in almost all tests. It is interesting to note that the box-plot representation somewhat conceals several significant weaknesses as detected when considering mean and standard deviation as done in Table I: For example, mean and standard deviation for the histogram bin variance are extremely high for Arnolds' cat map and Bakers' map (true also for sequence test values) while the values in the box-plots are not that bad. The reason are outliers with extremely poor values in case the interplay among iteration counter and other parameters leads to encrypted images close to the original by chance (which is a well known phenomenon for this type of chaotic ciphers). Those outliers of course affect mean and standard deviation more significantly. Finally we would like to emphasize the importance of considering large size datasets and keyspaces in experimentation underpinned by the large standard deviations

in the results of poor quality ciphers (Arnolds' cat map and Bakers' map in most cases). Employing a small dataset might therefore conceal such weaknesses.

9) Discussion on Security Tests Findings: The considered tests analyze the images' ciphertexts with respect to the distribution of the gray values. Only severe distribution defects of encrypted image data can be detected and only in case sufficiently large data are employed in such experiments. Furthermore, these tests do not (and can not) take into account the way how ciphertexts are generated. Algorithmic weaknesses in the encryption scheme, which may result in slight biases in the ciphertext, will likely not be revealed by tests that only look at the distribution of ciphertexts.

Some authors try to mimick advanced cryptanalytic techniques, such as differential cryptanalysis, and re-define them in a way that involves measurements on ciphertexts only. For example, computing the NCPR or UACI between two ciphertexts generated by two "similar" keys, is insufficient to evaluate security against differential attacks. A proper use of differential cryptanalysis usually requires the construction of vast amounts of ciphertexts whose plaintexts are related; furthermore, one often has to focus on parts of the encryption algorithm only (such as the output after a few rounds instead of the full cipher or a single S-box of the cipher) in order to uncover slight biases, which may be indicative of a security problem. The same applies to linear cryptanalysis, for which authors did not even try to provide a simplified version based on empirical tests discussed in Section III.

We thus postulate that the security of chaotic ciphers *cannot* be assessed thoroughly by any test that operates in a similar manner as those discussed in Section III (no consideration of the encryption algorithm, use of ciphertext properties only, unsystematic generation of test cases, no clear statistical decision criterion). In turn, any security analysis using these tests should raise suspicion.

Instead, the multimedia security community should apply the same techniques to analyze new ciphers, as routinely done in cryptography. Unfortunately, the IND-CPA definition yields no simple test criterion for security; indeed most symmetric ciphers (including AES) cannot formally be proven IND-CPA secure. Instead, one gains confidence in the security of a symmetric cipher by testing its resistance against all known attacks (including linear and differential cryptanalysis). As mentioned above, this requires a careful statistical analysis of the inner operations of the cipher under test. The works of [6], [25] are good examples: they scrutinize the chaos-based cipher at hand, focussing first on variants which contain only a few rounds, finally extending the results to the full cipher.

B. Encryption Speed

In order to assess the claim that the use of chaotic-based ciphers requires less computing resources than classic algorithms, we investigated the time that is required to encrypt an 512×512 image for various ciphers considered in this paper; Figure 6 depicts the results (milliseconds on the y-axis). All results are averaged over 128 images and 10 different keys. Experiments were conducted on a Linux machine with

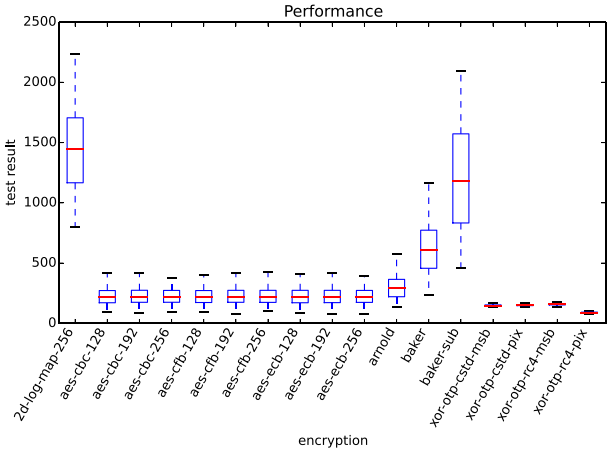

Fig. 6. Encryption speed: Time in milliseconds spent to encrypt an image.

4 AMD Opteron Processor 6174 processors with 12 cores, each 2,2 GHz, and 256 GB of main memory. Since our framework is single threaded, the number of cores is not relevant.

For these experiments, we use the encryption software already mentioned in Section III which is implemented in C++, open source and is freely available at GitHub.[8]

Arnold and Baker do not at all lead to good results. We can see that Baker without substitution is faster than with substitution, which is exactly the behaviour we would expect. The 2-D logistic map algorithm shows the highest computational cost overall.

Additionally, we compare the results to classic encryption algorithms that use AES in different modes. For the implementation of AES we use crypto++,[9] which is a free C++ cryptographic library. The naming convention for the AES modes is first *aes*, then the used mode (*cbc*, *cfb*, and *ecb*) and finally the used key length. As we can see, the AES approaches are significantly faster than all other considered chaos-based encryption algorithms.

From the numbers it is obvious that classic encryption is at least as competitive as chaos-based image encryption, in our results AES is even much faster. Due to the availability of highly tuned cryptographic libraries, it is hard for chaotic image encryption implementations to even reach the speed of fine-tuned classic ciphers. This questions the first main motivation, namely reduction of complexity. However, it has to be noted that eventually, when using highly tuned chaos-based encryption, this relation might change. In any case, the comparison has to be done against a tuned classical crypto library, which has never been done in literature on chaos-based encryption so far.

V. Conclusion

We have questioned recent developments in the area of chaos-based image and video encryption in two ways. First, we demonstrated that commonly used motivations to employ

[8]https://github.com/mpreis/seth

[9]http://www.cryptopp.com

these encryption primitives are not valid ones. Encrypting visual data with classic ciphers turns out to be significantly faster than chaos-based encryption implementation variants.

Furthermore, security concerns when applying classic strong block ciphers to redundant and correlated visual data do not apply in case these ciphers are used in the right way.

Second, we were able to demonstrate that deliberately chosen insecure encryption schemes (i.e., stream ciphers where the pseudorandom stream can be predicted) do either pass and/or perform very well in a battery of tests for experimental security evaluation, which are commonly used to assess chaos-based encryption schemes for visual data. Therefore, these metrics are clearly not usable for a sound evaluation of image ciphers – passing them is merely a necessary condition for security, but is by no means a sufficient criterion. This result fundamentally calls into question the security analysis of most prior works on chaotic image encryption.

Thus, in our view, authors of any publication that proposes a new image encryption scheme should:

- Choose a venue for submission (conference or journal) where security mechanisms and encryption schemes are in the core focus;
- Justify in a valid manner which advantages a new image encryption scheme brings over using a conventional strong cipher: Note that we have shown that there are no security concerns when using classical encryption techniques for visual data using proper techniques; additionally, the argument of superior computational performance of chaos-based schemes needs to be proven against fine-tuned crypto libraries (but not against, for example, hand-woven Matlab implementations); this requires the development of fine-tuned implementations of chaos-based schemes, preferably available as open source for the sake of results reproducibility; note also that eventually, superior computational performance as compared to highly tuned stream ciphers (e.g., those defined in the eStream portfolio) needs to be demonstrated, as these can also be applied to visual data without security concerns;
- Carefully analyze the available literature describing security breaches of algorithms of the same type/class (e.g., Section II) and explain, why the novel approach is not affected by existing cryptanalysis approaches;
- Analyze the security of a chaos-based cipher using methods and tools from cryptography and show that common cryptanalytic attacks (such as differential attacks) do not work against the new cipher. This requires scrutinizing the internal workings of the new encryption algorithm. Refrain from "proving" security by evaluating metrics on the ciphertexts only.

REFERENCES

[1] V. I. Arnold and A. Avez, *Problèmes Ergodiques De La Mécanique Classique*. Paris, France: Gauthier-Villars, 1967.

[2] J. Scharinger and F. Pichler, "Efficient image encryption based on chaotic maps," in *Proc. 20th Workshop Austrian Assoc. Pattern Recognit. (OAGM/AAPR) Pattern Recognit.*, Oldenbourg Verlag, Munich, Germany, 1996, pp. 159–170.

[3] J. Balatoni and A. Renji, "On the notion of entropy (Hungarian)," *Publ. Math. Inst. Hungarian Acad. Sci.*, vol. 1, no. 9, pp. 9–40, 1956.

[4] J. Fridrich, "Image encryption based on chaotic maps," in *Proc. 1997 IEEE Int. Conf. Syst., Man, Cybern., Comput. Cybern. Simulation*, Oct. 1997, pp. 1105–1110.

[5] J. Fridrich, "Symmetric ciphers based on two-dimensional chaotic maps," *Int. J. Bifurcation Chaos*, vol. 8, no. 6, p. 1259, 1998.

[6] L. Y. Zhang, C. Li, K.-W. Wong, S. Shu, and G. Chen, "Cryptanalyzing a chaos-based image encryption algorithm using alternate structure," *J. Syst. Softw.*, vol. 85, no. 9, pp. 2077–2085, 2012.

[7] X. Wang and K. Guo, "A new image alternate encryption algorithm based on chaotic map," *Nonlinear Dyn.*, vol. 76, no. 4, pp. 1943–1950, 2014.

[8] X. Wang, L. Teng, and X. Qin, "A novel colour image encryption algorithm based on chaos," *Signal Process.*, vol. 92, no. 4, pp. 1101–1108, 2012.

[9] H. Liu and Y. Liu, "Cryptanalyzing an image encryption scheme based on hybrid chaotic system and cyclic elliptic curve," *Opt. Laser Technol.*, vol. 56, pp. 15–19, Mar. 2014.

[10] A. Kanso and M. Ghebleh, "A novel image encryption algorithm based on a 3D chaotic map," *Commun. Nonlinear Sci. Numer. Simul.*, vol. 17, no. 7, pp. 2943–2959, 2012.

[11] A. A. A. El-Latif and X. Niu, "A hybrid chaotic system and cyclic elliptic curve for image encryption," *AEU—Int. J. Electron. Commun.*, vol. 67, no. 2, pp. 136–143, 2013.

[12] B. Furht, E. Muharemagic, and D. Socek, *Multimedia Encryption and Watermarking* (Multimedia Systems and Applications), vol. 28. New York, NY, USA: Springer, 2005.

[13] A. Uhl and A. Pommer, *Image and Video Encryption: From Digital Rights Management to Secured Personal Communication* (Advances in Information Security), vol. 15. New York, NY, USA: Springer, 2005.

[14] S. Lian, *Multimedia Content Encryption: Techniques and Applications*. Boca Raton, FL, USA: CRC Press, 2008.

[15] Y. Wu, J. P. Noonan, and S. Agaian, "NPCR and UACI randomness tests for image encryption," *Cyber J., Multidiscipl. J. Sel. Areas Telecommun.*, vol. 4, no. 2, pp. 31–38, 2011.

[16] G. Álvarez, F. Montoya, M. Romera, and G. Pastor, "Cryptanalyzing a discrete-time chaos synchronization secure communication system," *Chaos, Solitons Fractals*, vol. 21, no. 3, pp. 689–694, 2004.

[17] S. Li and X. Zheng, "Cryptanalysis of a chaotic image encryption method," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, May 2002, pp. II-708–II-711.

[18] S. Lian, J. Sun, and Z. Wang, "Security analysis of a chaos-based image encryption algorithm," *Phys. A, Statist. Mech. Appl.*, vol. 351, nos. 2–4, pp. 645–661, 2005.

[19] S. Li and X. Zheng, "On the security of an image encryption method," in *Proc. Int. Conf. Image Process.*, Sep. 2002, pp. II-925–II-928.

[20] C. Li, S. Li, K.-T. Lo, and K. Kyamakya, "A differential cryptanalysis of Yen–Chen–Wu multimedia cryptography system," *J. Syst. Softw.*, vol. 83, no. 8, pp. 1443–1452, 2010.

[21] S. Li, C. Li, G. Chen, and K.-T. Lo, "Cryptanalysis of the RCES/RSES image encryption scheme," *J. Syst. Softw.*, vol. 81, no. 7, pp. 1130–1143, 2008.

[22] J.-C. Yen and J.-I. Guo, "A new image encryption algorithm and its VLSI architecture," in *Proc. IEEE Workshop Signal Process. Syst.*, Oct. 1999, pp. 430–437.

[23] J.-C. Yen, H.-C. Chen, and S.-M. Wu, "Design and implementation of a new cryptographic system for multimedia transmission," in *Proc. IEEE Int. Symp. Circuits Syst.*, May 2005, pp. 6126–6129.

[24] A. Rukhin, J. Soto, J. Nechvatal, M. Smid, and E. Barker, "A statistical test suite for random and pseudorandom number generators for cryptographic applications," Nat. Inst. Standard Technol., Gaithersburg, MD, USA, NIST Special Pub. 800-22, Revision 1a, 2010.

[25] W.-S. Yap, R. C.-W. Phan, W.-C. Yau, and S.-H. Heng, "Cryptanalysis of a new image alternate encryption algorithm based on chaotic map," *Nonlinear Dyn.*, vol. 80, no. 3, pp. 1483–1491, 2015.

[26] Y. Zhang, Y. Wang, and X. Shen, "A chaos-based image encryption algorithm using alternate structure," *Sci. China F, Inf. Sci.*, vol. 50, no. 3, pp. 334–341, 2007.

[27] H.-C. Chen, J.-C. Yen, and J.-I. Guo, "Design of a new cryptography system," in *Advances in Multimedia Information Processing—PCM* (Lecture Notes in Computer Science), vol. 2532. Berlin, Germany: Springer-Verlag, 2002, pp. 1041–1048.

[28] H.-C. Chen and J.-C. Yen, "A new cryptography system and its VLSI realization," *J. Syst. Archit.*, vol. 49, nos. 7–9, pp. 355–367, 2003.

[29] M. Feki, B. Robert, G. Gelle, and M. Colas, "Secure digital communication using discrete-time chaos synchronization," *Chaos, Solitons Fractals*, vol. 18, no. 4, pp. 881–890, 2003.
[30] J. C. Yeo and J. I. Guo, "Efficient hierarchical chaotic image encryption algorithm and its VLSI realisation," *IEE Proc.-Vis., Image Signal Process.*, vol. 147, no. 2, pp. 167–175, Apr. 2000.
[31] C. Li, "Cracking a hierarchical chaotic image encryption algorithm based on permutation," *Signal Process.*, vol. 118, pp. 203–210, Jan. 2016.
[32] C. Li, L. Y. Zhang, R. Ou, K.-W. Wong, and S. Shu, "Breaking a novel colour image encryption algorithm based on chaos," *Nonlinear Dyn.*, vol. 70, no. 4, pp. 2383–2388, 2012.
[33] C. Li, Y. Liu, T. Xie, and M. Z. Q. Chen, "Breaking a novel image encryption scheme based on improved hyperchaotic sequences," *Nonlinear Dyn.*, vol. 73, no. 3, pp. 2083–2089, 2013.
[34] N. K. Pareek, V. Patidar, and K. K. Sud, "Image encryption using chaotic logistic map," *Image Vis. Comput.*, vol. 24, no. 9, pp. 926–934, 2006.
[35] C. Li, S. Li, M. Asim, J. Nunez, G. Alvarez, and G. Chen, "On the security defects of an image encryption scheme," *Image Vis. Comput.*, vol. 27, no. 9, pp. 1371–1381, 2009.
[36] Z.-H. Guan, F. Huang, and W. Guan, "Chaos-based image encryption algorithm," *Phys. Lett. A*, vol. 346, nos. 1–3, pp. 153–157, 2005.
[37] C. Çokal and E. Solak, "Cryptanalysis of a chaos-based image encryption algorithm," *Phys. Lett. A*, vol. 373, no. 15, pp. 1357–1360, 2009.
[38] E. Solak, C. Çokal, O. Yildiz, and T. Biyikoğlu, "Cryptanalysis of Fridrich's chaotic image encryption," *Int. J. Bifurcation Chaos*, vol. 20, no. 5, pp. 1405–1413, 2010.
[39] E. Y. Xie, C. Li, S. Yu, and J. Lü, "On the cryptanalysis of Fridrich's chaotic image encryption scheme," *Signal Process.*, vol. 132, pp. 150–154, Mar. 2017.
[40] G. Alvarez, J. M. Amigó, D. Arroyo, and S. Li, "Lessons learnt from the cryptanalysis of chaos-based ciphers," in *Chaos-Based Cryptography: Theory, Algorithms and Applications*. Berlin, Germany: Springer, 2011, pp. 257–295.
[41] Y. Wu, Y. Zhou, G. Saveriades, S. Agaian, J. P. Noonan, and P. Natarajan, "Local Shannon entropy measure with statistical tests for image randomness," *Inf. Sci.*, vol. 222, pp. 323–342, Feb. 2013. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S002002551200521X
[42] S. Jenisch and A. Uhl, "Visual security evaluation based on SIFT object recognition," in *Artificial Intelligence Applications and Innovations*. Berlin, Germany: Springer-Verlag, 2014, pp. 624–633.
[43] Y. Wu, J. P. Noonan, G. Yang, and H. Jin, "Image encryption using the two-dimensional logistic chaotic map," *J. Electron. Imag.*, vol. 21, no. 1, p. 013014, 2012.
[44] V. Patidar, N. K. Pareek, and K. Sud, "A new substitution–diffusion based image cipher using chaotic standard and logistic maps," *Commun. Nonlinear Sci. Numer. Simul.*, vol. 14, no. 7, pp. 3056–3075, 2009.
[45] C. Li, T. Xie, Q. Liu, and G. Chen, "Cryptanalyzing image encryption using chaotic logistic map," *Nonlinear Dyn.*, vol. 78, no. 2, pp. 1545–1551, 2014.
[46] Z. Hua, Y. Zhou, C.-M. Pun, and C. L. P. Chen, "2D sine logistic modulation map for image encryption," *Inf. Sci.*, vol. 297, pp. 80–94, Mar. 2015.
[47] Y. Zhou, L. Bao, and C. Chen, "Image encryption using a new parametric switching chaotic system," *Signal Process.*, vol. 93, no. 11, pp. 3039–3052, 2013.
[48] S. Fluhrer, I. Mantin, and A. Shamir, "Weaknesses in the key scheduling algorithm of RC4," in *Selected Areas in Cryptography* (Lecture Notes in Computer Science), vol. 2259. Berlin, Germany: Springer-Verlag, 2001, pp. 1–24.
[49] R. Ye and W. Guo, "A chaos-based image encryption scheme using multimodal skew tent maps," *J. Emerg. Trends Comput. Inf. Sci.*, vol. 4, no. 10, pp. 800–810, 2013.
[50] G. Yu, Y. Shen, G. Zhang, and Y. Yang, "A chaos-based color image encryption algorithm," in *Proc. 6th Int. Symp. Comput. Intell. Design (ISCID)*, Oct. 2013, pp. 92–95.
[51] Z. Yu, Z. Zhe, Y. Haibing, P. Wenjie, and Z. Yunpeng, "A chaos-based image encryption algorithm using wavelet transform," in *Proc. 2nd Int. Conf. Adv. Comput. Control*, Mar. 2010, pp. 217–222.
[52] F. Wu, W. Cui, and H. Chen, "A compound chaos-based encryption algorithm for vector geographic data under network circumstance," in *Proc. Congr. Image Signal Process. (CISP)*, May 2008, pp. 254–258.
[53] B. Boulebtateche, M. M. Lafifi, and S. Bensaoula, "A multimedia chaos-based encryption algorithm," in *Proc. 12th Int. Arab Conf. Inf. Technol. (ACIT)*, Riyadh, Saudi Arabia, Dec. 2011.
[54] T. Gao and Z. Chen, "A new image encryption algorithm based on hyper-chaos," *Phys. Lett. A*, vol. 372, no. 4, pp. 394–400, 2008.
[55] X. Wu, "A novel chaos-based image encryption scheme using coupled map lattices," in *Proc. 10th Int. Conf. Fuzzy Syst. Knowl. Discovery*, Jul. 2013, pp. 1020–1024.
[56] L. Zhang, X. Liao, and X. Wang, "An image encryption approach based on chaotic maps," *Chaos, Solitons Fractals*, vol. 24, no. 3, pp. 759–765, 2005.
[57] M. Ahmad and H. D. Al-Sharari. (2014). "An inter-component pixels permutation based color image encryption using hyper-chaos." [Online]. Available: https://arxiv.org/abs/1403.4780
[58] A. S. Alghamdi, H. Ullah, M. Mahmud, and M. K. Khan, "Bio-chaotic stream cipher-based iris image encryption," in *Proc. Int. Conf. Comput. Sci. Eng.*, Aug. 2009, pp. 739–744.
[59] Y. Cao, R. Qiu, and Y. Fu, "Color image encryption based on hyper-chaos," in *Proc. 2nd Int. Congr. Image Signal Process. (CISP)*, Oct. 2009, pp. 1–6.
[60] M. Farajallah, Z. Fawaz, S. El Assad, and O. Deforges, "Efficient image encryption and authentication scheme based on chaotic sequences," in *Proc. 7th Int. Conf. Emerg. Secur. Inf., Syst. Technol.*, 2013, pp. 150–155.
[61] H. Khanzadi, M. Eshghi, and S. E. Borujeni, "Image encryption using random bit sequence based on chaotic maps," *Arabian J. Sci. Eng.*, vol. 39, no. 2, pp. 1039–1047, 2014.
[62] Q. Run-he, C. Yun, and F. Yu-Zhen, "Integrated confusion-diffusion mechanisms for chaos based image encryption," in *Proc. 4th Int. Congr. Image Signal Process.*, Oct. 2011, pp. 629–632.
[63] E. Alvarez, A. Fernández, P. García, J. Jiménez, and A. Marcanoc, "New approach to chaotic encryption," *Phys. Lett. A*, vol. 263, nos. 4–6, pp. 373–375, 1999.
[64] L. Li-Hong, B. Feng-Ming, and H. Xue-Hui, "New image encryption algorithm based on logistic map and hyper-chaos," in *Proc. 5th Int. Conf. Comput. Inf. Sci.*, Jun. 2013, pp. 713–716.
[65] A. Awad and D. Awad, "Efficient image chaotic encryption algorithm with no propagation error," *ETRI J.*, vol. 32, no. 5, pp. 774–783, 2010.
[66] M. François, T. Grosges, D. Barchiesi, and R. Erra, "Image encryption algorithm based on a chaotic iterative process," *Appl. Math.*, vol. 3, no. 12, pp. 1910–1920, 2012.
[67] A. G. Radwan, S. H. AbdElHaleem, and S. K. Abd-El-Hafiz, "Symmetric encryption algorithms using chaotic and non-chaotic generators: A review," *J. Adv. Res.*, vol. 7, no. 2, pp. 193–208, 2016.
[68] H.-I. Hsiao and J. Lee, "Color image encryption using chaotic nonlinear adaptive filter," *Signal Process.*, vol. 117, pp. 281–309, Dec. 2015.
[69] B. Stoyanov and K. Kordov, "Image encryption using chebyshev map and rotation equation," *Entropy*, vol. 17, no. 4, pp. 2117–2139, 2015.

Mario Preishuber received the bachelor's degree from the University of Salzburg in 2014, where he is currently pursuing the master's degree. He is working on his master thesis which is about effective memory reuse, supervised by Prof. C. Kirsch. In 2017, he was invited to Google's 5th Compiler and Programming Language Summit 2017. He was a Visiting Student at the École Polytechnique Fédérale de Lausanne, Switzerland, during 2015/2016. During the last years, he was project staff at academic projects in collaboration with research groups of Prof. A. Uhl and Prof. C. Kirsch.

Thomas Hütter received the bachelor's and master's degrees in computer science from the University of Salzburg, Austria, in 2014 and 2017, respectively, where he is currently pursuing the Ph.D. degree with the Database Research Group, under the supervision of Prof. N. Augsten. His research interests include efficient algorithms for complex data structures and similarity search in massive data collections.

Stefan Katzenbeisser (S'98–A'01–M'07–SM'12) received the Ph.D. degree from the Vienna University of Technology, Austria. After he was a Research Scientist with the Technical University of Munich, Germany, he joined Philips Research as a Senior Scientist in 2006. Since 2008, he has been a Professor with the Technische Universität Darmstadt, heading the Security Engineering Group. He has authored over 200 scientific publications and served on the program committees of several workshops and conferences devoted to information security. His current research interests include embedded security, data privacy, and cryptographic protocol design. He is currently serving on the Information Forensics and Security Technical Committee of the IEEE Signal Processing Society.

Andreas Uhl received the Ph.D. degree from the University of Salzburg. He is currently a Professor with the Department of Computer Science, University of Salzburg. He has co-authored over 400 scientific publications. His research interests are in processing and analysis of visual data in general, and in biometric systems, multimedia security and forensics, and medical data analysis in particular. He acts as an Associate Editor for ACM TOMM, *Signal Processing: Image Communication*, the *Journal of Visual Communication and Image Representation*, and *ETRI Journal*.