Proceedings of the European Control Conference 2007
Kos, Greece, July 2-5, 2007



# A GENERALIZATION OF BENFORD'S LAW AND ITS APPLICATION TO IMAGES

*Fernando Pérez-González* †, *Gregory L. Heileman* ∗ *and Chaouki T. Abdallah* ∗

† Dept. Teoría de la Señal y Comunicaciones, ETSI Telecom., Universidad de Vigo, 36200 Vigo, Spain;
∗ ECE Dept. University of New Mexico, Albuquerque, NM 87131, USA

## ABSTRACT

We present a generalization of Benford's law for the first significant digit. This generalization is based on keeping two terms of the Fourier expansion of the probability density function of the data in the modular logarithmic domain. We prove that images in the Discrete Cosine Transform domain closely follow this generalization. We use this property to propose an application in image forensics, namely, detecting that a given image carries a hidden message.

***Index Terms***— Benford's law, DCT, Fourier series, forensics, watermarking.

## 1. INTRODUCTION

Benford's law of "anomalous digits" was enunciated by General Electric's physicist Frank L. Benford in 1938 [1], and predicts the frequency of appearance of the most significant digit (MSD) for a broad range of natural and artificial data. Even though there exists a generalization of Benford's law for digits in other significant places, in this paper we will only deal with the MSD as the general principles that we will discuss throughout directly derive from such case. Given a number in decimal form, the MSD is simply the leading digit of the mantissa (assuming that the exponent is a power of 10); hence, the MSD cannot take the value 0. For instance, the MSD of 2.85 is 2, and the MSD of 0.0034 is 3. Benford's law specifies that the probability that the MSD take the value $d \in \{1, 2, \cdots, 9\}$ is $P(d) = \log_{10}(1 + 1/d)$. Benford's original paper does not furnish any theoretical explanation for the law, but simply verifies it empirically on various data, ranging from lengths of river waterbeds to population figures. In fact, Benford seemingly arrived at his law by noticing that the pages in logarithm books were unevenly worn out. Quite strikingly, exactly the same law had been enunciated more than fifty years earlier by S. Newcomb [2] with logarithm tables also acting as catalyst.

Since the publication of Benford's paper, many researchers have made significant contributions at both the fundamental and the application levels. It can be safely said that the underlying mechanisms that make Benford's law hold in many useful situations are known; these will be briefly reviewed in Section 2. On the other hand, at a practical level, Benford's law has been shown to apply to half-life time of radioactive particles [3], financial data [4], regression coefficients [5], dynamical systems [6] and many others. Of particular interest for our purposes is the work by J.M. Jolion [7], who showed that Benford's law holds reasonably well in gradient images and in pyramidal decompositions based on the Laplace transform. To the best of our knowledge, the only other work dealing with Benford's law for images is due to E. Acebo and M. Sbert [8], who proposed the use of Benford's law to determine whether synthetic images were generated using physically realistic methods, although the fact that many real images do not follow Benford's law (see Section 3) puts this application in question.

In this paper we show that while images in the "pixel" domain seem not to obey Benford's law, the situation changes quite dramatically when they are transformed using the Discrete Cosine Transform (DCT). Furthermore, we will present a generalization of Benford's law, based on Fourier analysis, that leads to a much closer fit to the observed digits frequencies. We will also give a theoretical explanation of why images in the DCT domain satisfy the generalized law; such explanation heavily relies on well known and thoroughly tested statistical properties of DCT coefficients. Finally, we will hint at some possible applications in forensics, by showing how the Fourier-based formulation can be used to detect whether an image has been watermarked.

## 2. BACKGROUND

In this Section we will recall some of the known properties that affect random variables in the context of Benford's law. The next proposition [6] helps us establish which random variables satisfy Benford's law

**Proposition 1** *A random variable $X$ follows Benford's law if the random variable $Y = \log_{10} X \mod 1$ is uniform in $[0, 1)$.*

The proof is based on the fact that the intervals corresponding to $d = k$ are of the form $[k, k + 1) \cdot 10^n$, for any integer $n$. Now, after taking the base-10 logarithm the intervals take the form $n + [\log_{10} k, \log_{10}(k+1))$, with $n$ integer. After modulo-1 reduction, all intervals fold onto $[\log_{10} k, \log_{10} k + 1)$. Integrating a uniform pdf over this interval we obtain $\log_{10}(1 + \frac{1}{k})$.

ISBN 978-3-9524173-8-6

Note that this condition is sufficient but not necessary: one may perfectly construct other non-uniform pdf's for which Benford's law is satisfied. Random variables satisfying the property in Prop. 1 are sometimes called *strong Benford*, a term to which we will adhere henceforth. Moreover, given a real number $x$, we will say that $\log_{10} x \mod 1$ is defined in the *Benford domain*. Returning to the proof of Prop. 1, notice that at a more intuitive level, the rationale behind Benford's law lies in the fact that, after transforming a random variable to the $\log_{10}$ domain, the width of the intervals assigned to each digit is different.

**Proposition 2 Scale invariance:** *Suppose that $X$ follows Benford's law; then the random variable $Z = \alpha X$ will follow Benford's law for an arbitrary $\alpha$ if and only if $X$ is strong Benford.*

The proof is as follows: consider the random variable $\log_{10} Z \mod 1 = \log_{10} \alpha + \log_{10} X \mod 1$. This corresponds to a cyclic shift of $\log_{10} \alpha \mod 1$ on the pdf of $Y = \log_{10} X \mod 1$. Now, for this shift to leave the same probabilities regardless of the value of $\alpha$ one must have that $Y$ is uniform in $[0, 1)$.

**Definition 1** *Let $X$ be a random variable following a certain law $\phi_{10}(d)$ for the MSD. Then, $X$ is base invariant if for any base of the form $10^n$, with $n \in \mathbb{N}$, then $\phi_{10^n}(d) = \phi_{10}(d)$, where $\phi_{10^n}(d)$ is the probability that the first digit of $X$ in a base $10^n$ representation is $d$.*

The next proposition, stated without proof concerns invariance to change of base. For a proof, see [9].

**Proposition 3 Base invariance:** *The random variable $X$ with continuous distribution is base invariant if and only if $X$ is strong Benford.*

Another very useful property concerns the product of random variables.

**Proposition 4 Product of independent random variables:** *Let $X$ be strong Benford, and let $Y$ be another random variable independent of $X$ and such that $Y \neq 0$. Then, the random variable $Z = X \cdot Y$ is strong Benford.*

For the sketch of a proof, notice that $\log_{10} Z = \log_{10} X + \log_{10} Y$. Therefore, the pdf of $\log_{10} Z$ will be the convolution of those of $\log_{10} X$ and $\log_{10} Y$. However, we are interested in the modulo-1 reduction of $\log_{10} Z$; it can be easily shown that this pdf can be obtained by performing the convolution between the pdf of $\tilde{X} \triangleq \log_{10} X \mod 1$ and that of $\log_{10} Y$, and reducing the result modulo-1. Alternatively, one can do the circular convolution over $[0, 1)$ of the pdf's of $\tilde{X}$ and $\tilde{Y} \triangleq \log_{10} Y \mod 1$. However, when $\tilde{X}$ is uniform in $[0, 1)$ the circular convolution is always uniform in $[0, 1)$, regardless of the distribution of $\tilde{Y}$.

**Fig. 1**. Figure 'Man' used in the experiments.

The previous property has important consequences, as the proof can be adapted to show that the product of $N$ independent random variables asymptotically tends to a random variable that obeys Benford's law [10]. Hence, if the data are drawn from the product of a number of random variables, the MSD of the samples will be close to Benford's law.

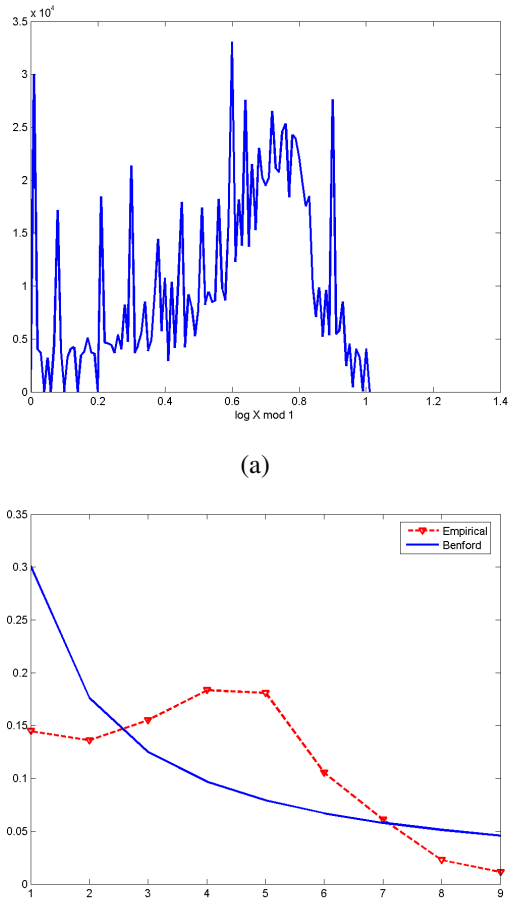
The product interpretation also connects Benford's law to *mixtures of random variables*. Mixtures of random variables are relevant in image processing after the proof by Hjorungnes et al. [11] that a Laplacian distribution (often used to model the coefficients of a block-wise DCT transform) can be written as a mixture of Gaussians whose variance is controlled by an exponential distribution. Thus, if $f_X(x)$ denotes a zero-mean unit-variance Gaussian pdf, the mixture takes the form

$$f_Z(z) = \int_0^\infty f_X(z|\sigma^2) f_{\sigma^2}(\sigma^2) d\sigma^2 \qquad (1)$$

where $f_{\sigma^2}(\sigma^2)$ is an exponential. Interestingly, mixtures of the general form given in (1) can be written in such a way that Proposition 4 can be straightforwardly applied. Indeed, the random variable $Z$ whose pdf is $f_Z(z)$ is obtained through (1) can be written as $Z = X \cdot \sqrt{\Sigma}$, with $\Sigma$ the random variable that controls the variance. From here, it is immediate to conclude that if either $X$ or $\sqrt{\Sigma}$ conform to Benford's law, then $Z$ will also do so. Additional insight can be obtained by seeing (1) as an iterator that takes certain distribution (of $X$) and outputs another distribution (of $Z$), but this will not pursued further.

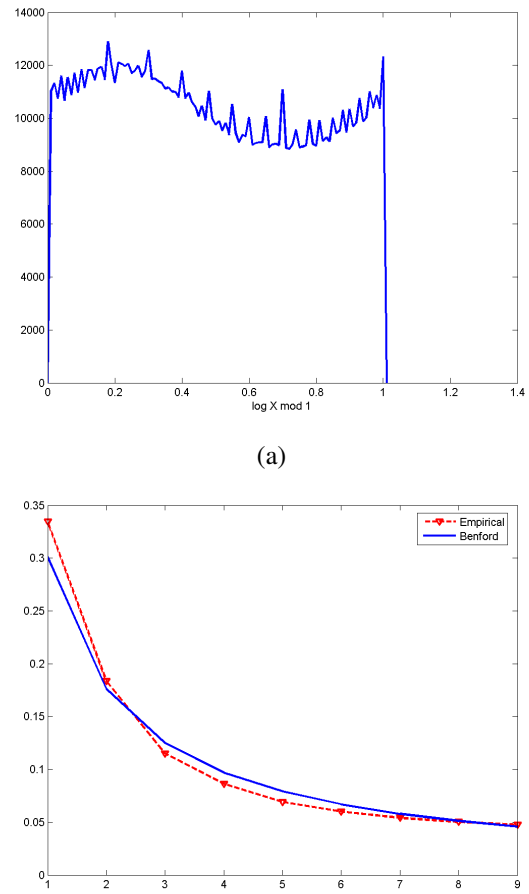
## 3. APPLICATION OF BENFORD'S LAW TO IMAGES

Given the seemingly good match of natural phenomena to Benford's law, it is reasonable to ask whether this will be so




(a) (b)

**Fig. 2**. Histogram of the luminance values of 'Man' in Benford ($\log_{10} \mod 1$) domain (a); Distribution of the MSD corresponding to 'Man' (b).




(a) (b)

**Fig. 3**. Histogram of the DCT values of 'Man' in Benford ($\log_{10} \mod 1$) domain (a); Distribution of the MSD corresponding to 'Man' (b). Block size is $8 \times 8$.

for images. Unfortunately, it is well known that image luminance possess a histogram that does not admit a closed-form, as there is a strong variation from picture to picture. Hence, it is highly unlikely that Benford's law, or a generalization, be applicable here. Our experiments confirm that gray-level images do not satisfy Benford's law. To illustrate, consider the image 'Man' shown in Fig. 1 for which the histogram of the variable $\log_{10} X \mod 1$ is shown in Fig. 2(a). Clearly this histogram falls short of being constant, which would guarantee compliance to Benford's law. Consequently, the MSD distribution is quite different from that proposed by Benford, as plotted in Fig. 2(b).

However, the results are quite different results when we consider the block-wise DCT transform, whose coefficients match Benford's law reasonably well. Figure 3(a) shows the histogram of the variable $\log_{10} X \mod 1$, with $X$ given in the block-DCT domain, for the image 'Man', while Fig. 3(b) represents the distribution of the MSD, which now lies closer to Benford's distribution. The DCT block size used in these figures is $8 \times 8$; however, similar results are obtained by considering other block sizes as well as other images. A crucial observation from Fig. 3(a) is that the histogram is not really flat, but instead can be modeled with a constant plus a sinusoidal (AC) term. This somehow surprising phenomenon was observed in all images we tried, thus suggesting a generalization of Benford's law to accommodate the extra term. The crucial question is therefore why DCT coefficients follow this generalized form of Benford's law. It is well known that the coefficients of a block-based DCT can be accurately modeled by a Generalized Gaussian (GG) distribution. For instance, let $b^{(n)}(i,k)$, $i,k \in \{0,\cdots,7\}$ denote the $(i,k)$-th coefficient of the DCT of the $n$-th block. Then, $b^{(n)}(i,k)$ for all $n$ can be thought of as being drawn from a GG distribution. A GG distribution has the form $f_X(x) = Ae^{-|\beta x|^c}$ where $A$

and $\beta$ are expressed in terms of $c$ and the standard deviation $\sigma$ as follows

$$\beta = \frac{1}{\sigma}\left(\frac{\Gamma(3/c)}{\Gamma(1/c)}\right)^{1/2}; \;\; A = \frac{\beta c}{2\Gamma(1/c)} \tag{2}$$

The parameter $c$ is also regarded to as the *shaping factor*. Unfortunately, the parameters $\sigma$ and $c$ vary with the frequency indices $(i,k)$. This implies that the coefficients that we are modeling should be rather considered as being generated by a mixture of GGs, in which the parameters are governed by a certain rule. This might complicate the derivation of a statistical model for the variables in the Benford domain; fortunately, in the next section we will see how this difficulty can be overcome by averaging with respect to the parameters.

## 4. A FOURIER-SERIES-BASED MODEL

The sinusoidal character of the histogram in Fig. 3(a) suggests that a Fourier representation for the pdf of the variable $\tilde{X} \triangleq \log_{10} X \mod 1$ would be plausible. For a further justification, let $X' \triangleq \log_{10} X$, then we can write

$$f_{\tilde{X}}(x) = \sum_{k=-\infty}^{\infty} f_{X'}(x-k) = f_{X'}(x) * \sum_{k=-\infty}^{\infty} \delta(x-k), x \in [0,1) \tag{3}$$

where $*$ denotes convolution and $\delta(x)$ is Dirac's delta. Taking the Fourier transform on both sides of the equation above, we can write

$$\begin{aligned} \Phi_{\tilde{X}}(\omega) &= \Phi_{X'}(\omega) \sum_{k=-\infty}^{\infty} \delta(\omega - 2\pi k) \\ &= \sum_{k=-\infty}^{\infty} \Phi_{X'}(2\pi k)\delta(\omega - 2\pi k) \end{aligned} \tag{4}$$

From here, it follows that we can write the pdf of the variable in the modular logarithmic domain as

$$f_{\tilde{X}}(x) = 1 + 2\sum_{k=1}^{\infty} |a_k| \cos(2\pi k x + \phi_k), x \in [0,1) \tag{5}$$

where $a_k = |a_k| e^{j\phi_k} \triangleq \Phi_{X'}(2\pi k)$.

Note that this representation is valid as long as the conditions for convergence of the Fourier series are met. However, the cases of specific interest to us are those for which the magnitude of the Fourier coefficients $|a_k|$ is small for moderate and large $k$. Note that the case of a strong Benford random variable corresponds to $|a_k| = 0$ for all $k \geq 1$.

We want to show that a GG random variable can be very accurately modeled in the Benford domain by a distribution composed of a constant and one sinusoidal term. To this end, we compute the coefficients of its Fourier series

$$\begin{aligned} a_n &= \int_{-\infty}^{\infty} f_{\tilde{X}}(x) e^{-j2\pi n x} dx \\ &= 2A \int_0^{\infty} \exp(-(\beta z)^c) \exp(-j2\pi n \log z / \log 10) dz \\ &= \frac{2A e^{j\frac{2\pi n \log \beta}{\log 10}}}{\beta} \int_0^{\infty} \exp(-(z)^c) z^{-j2\pi n/\log 10} dz \\ &= \frac{2A}{\beta c} e^{j\frac{2\pi n \log \beta}{\log 10}} \Gamma\left(\frac{-j2\pi n + \log 10}{c \log 10}\right) \end{aligned} \tag{6}$$

Several interesting comments are due:

- The absolute value of the coefficients depends only on $A/\beta$, and hence, is independent of the variance.
- The only effect of the variance is a phase change. Suppose that a certain phase shift is achieved for $\beta'$; then, it is easy to see that the same shift will be achieved for any $\beta = 10^k \beta'$, with $k$ integer.

However, the most important observation is that the magnitude of the coefficients in (6) decreases very rapidly with $n$. In fact, using [12, Eq. 8.326] it is immediate to show that

$$|a_n|^2 = \prod_{k=0}^{\infty} \left(1 + \frac{(2\pi n)^2}{\log^2(10)(ck+1)^2}\right)^{-1} \tag{7}$$

which, if truncated, quickly converges to the true value. Moreover, from (7) it is easy to see that the Fourier series coefficients monotonically increase with the shaping factor $c$.

To get an idea of the magnitude of $a_n$, we evaluated (7) for different values of $c$. For a Gaussian (i.e., $c = 2$) the following magnitudes are obtained while for a Laplacian (i.e., $c = 1$) we

| $n$ | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| $\lvert a_n \rvert$ | 0.165849 | 0.0194532 | 0.00228155 | 0.00026759 |

get the following values Finally, for $c = 0.5$ we observe that

| $n$ | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| $\lvert a_n \rvert$ | 0.0569 | 0.00110 | $1.866 \cdot 10^{-5}$ | $2.964 \cdot 10^{-7}$ |

the magnitudes are so small that even $|a_1| = 0.00614761$.

The main consequence of all these evaluations is that for all values of the shaping gain smaller than 2 (which are typical in images), the following approximation

$$f_{\tilde{X}}(x) \approx 1 + 2|a_1| \cos(2\pi k x + \phi_k), x \in [0,1) \tag{8}$$

is reasonable. However, unless $c$ is small the strong Benford property does not hold, so one must keep the term corresponding to the first coefficient of the Fourier series expansion in order to ensure a small approximation error. In Fig. 4 we plot the theoretical MSD distribution that results after using the

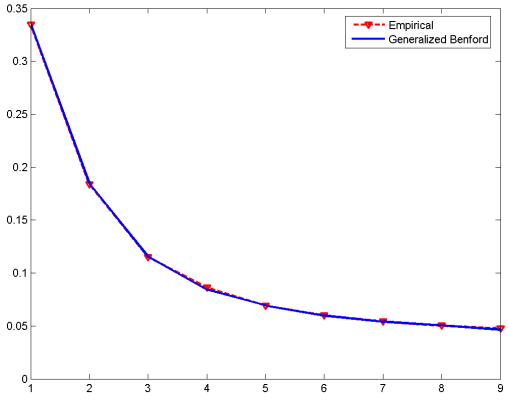



**Fig. 4**. Empirical digit distribution and generalized Benford's law. $a_1 = 0.067$; $\phi_1 = -1.221$ rad.

approximation in (8), which leads to an excellent agreement with the empirical distribution. We note that although the Fourier coefficient $a_1$ can be obtained by projecting the histogram of $\log_{10} X$ onto the complex exponential $\exp(j2\pi x)$, a simpler procedure is possible through the use of sampled complex moments.

So far, we have shown that a GG distribution can be closely approximated as in (8). However, as we have remarked, different DCT coefficients will have different parameters $\sigma$ and $c$. Suppose that these two parameters can be modeled as being drawn from a joint distribution $f_{C,\Sigma}(c,\sigma)$. Then, the pdf of the variable in the modular logarithmic domain can be written as

$$\begin{aligned} f_{\bar{X}}(x) &= 1 + \int\int \sum_{\substack{k=-\infty \\ k\neq 0}}^{\infty} a_k(c,\sigma) e^{j2\pi kx} f_{C,\Sigma}(c,\sigma) dc \cdot d\sigma \\ &= 1 + 2\mathrm{Re}\{\sum_{k=1}^{\infty} \overline{a}_k e^{j2\pi kx}\} \end{aligned} \quad (9)$$

where $\overline{a}_k$ is the mean value of $a_k$ averaged over the joint distribution of $C$ and $\Sigma$. Then, as long as the averaged coefficients $\overline{a}_k$ are such that their magnitude is small for $k > 1$, the approximation of the form (8) is valid.

Now suppose that $c$ is such that for all $i, k$, $c(i,k) \leq c^+$, then the following can be proven for all $(i,k)$ and all $n$

$$|a_n(c,\sigma)| \leq |a_n(c^+)| \quad (10)$$

again suggesting that for values of $c^+$ less than 2, as is customary in practice, the approximation given in (8) is valid.

The previous discussion has important implications in video: if all frames of a video sequence can be modeled as in (8), then the whole sequence will also satisfy this property. Therefore, our generalized form of Benford's law will apply to video sequences as well, provided that one works with the block-DCT coefficients of each frame.

Let us remark that all our results concerning Generalized Gaussian distributions apply to zero-mean random variables. It is not difficult to see and simulate that as the mean is increased, the distribution of the base-10 digits gets away from Benford's distribution.

## 5. IMAGE FORENSICS IN THE BENFORD DOMAIN

As pointed out in the Introduction, Benford's law has been successfully applied to detect fraud in tax data [13], [4]. The test is based on the assumption that real data follow Benford's law on the basis that they come from many independent sources with different scales (much the same as data in a newspaper, which also approximately satisfies the law). Another application in forensics is to detect data manipulation in scientific data: Diekmann [5] showed that the first digits of regression coefficients from sociological analyses closely approximate the Benford distribution. Diekmann went on to conduct an experiment in which he asked students in a statistics course to fabricate regression coefficients; while the hypothesis of the MSD from the fabricated data not following Benford's law could not be rejected, this was not the case when the second most significant digit was considered, with a statistically significant difference. Schäfer *et al.* [14] investigated the use of Benford's law to analyze data from surveys, where there is always the risk that the interviewers fabricate the data. They focused on data from the German Socio-Economic Panel which contained proven fakes, detected after a second wave. Once again, Benford's law served to spot those interviewers who had cheated: by analyzing the full set of answers to specific questions known to conform to Benford's law (such as net-income or tax data), it was possible to easily detect fabrications.

In view of the above applications, it is natural to ask whether Benford's law may find any use in image forensics. Although we foresee other specific applications which are discussed in the Conclusions, here we focus on *image steganography*. In steganography one is interested in hiding information in such a way that it remains undetected. In fact, the purpose of the attacker (steganalyst) is to detect the presence of the hidden message, leaving to higher layers the attempts to extract it. Steganography is closely related to *watermarking*; the main difference is that while robustness is the major concern for the latter, detectability is the key issue for the former. There is a vast amount of algorithms for image steganography or watermarking; likewise there are a number of algorithms for steganalysis, the most sophisticated ones paying attention to marginal or joint histograms of the coefficients, or related statistics. For instance, Lyu and Farid [15] have proposed the use of higher order statistics in steganalysis.

We have seen in the previous section how DCT coefficients of natural images conform to a generalized form of Benford's law for which the AC coefficients of the Fourier expansion in the Benford domain are very small, except for

the first one (i.e., $|a_1|$). Then, it is reasonable to think of using the following test:

1. Compute the magnitudes of the Fourier coefficients in the Benford domain.
2. Determine the "noise-level" by averaging the Fourier coefficients over a suitable window. This noise-level consists of a mean and a variance.
3. Find the index $n_*$ of the coefficient such that its magnitude is greater (in a statistically significant sense) than the noise-level.
4. If $n_* > 1$, then the image is declared as watermarked.

To verify the plausibility of the proposed test, we have watermarked the 'Man' image in the DCT domain using the spread-spectrum method proposed and analyzed in [16]. This method allows to achieve a better concealment of the watermark by taking into account the characteristics of the human visual system through the computation of a so-called *perceptual mask* that modulates the amplitude of the watermark. It is important to remark that the method considered here somewhat favors a Benford-inspired steganalysis, because watermarking takes place in the same domain as the data that conforms to the generalized Benford distribution. Watermarking in other domains (e.g., the spatial) might be harder to detect with the proposed test, nevertheless exhaustive experiments have yet to be carried out.

Table 1 summarizes the results obtained for the original image 'Man' (non watermarked). For each value of $n$, we represent the value of $|a_n|$, and the mean $\mu_n$ and typical deviation $\sigma_n$ of the vector $(|a_{n+1}|, \cdots, |a_{n+L}|)^T$, where $L$ is the window-length (in our experiments set to 8) and $T$ denotes transpose. Again, we select the largest index $n_*$ such that $|a_{n_*}| > \mu_{n_*} + 2.58\sigma_{n_*}$. This value of $2.58\sigma_n$ guarantees a probability of false positives less than 0.01 under a Gaussian distribution.

| $n$ | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| $\|a_n\|$ | 0.0670 | 0.0032 | 0.0035 | 0.0037 | 0.0015 |
| $\mu_n$ | 0.0032 | 0.0029 | 0.0027 | 0.0028 | 0.0028 |
| $\sigma_n$ | 0.0013 | 0.0016 | 0.0016 | 0.0017 | 0.0017 |

**Table 1**. Fourier coefficients and noise-level parameters for the original image.

Clearly, for this case $n_* = 1$, thus suggesting that the image under study had not been watermarked.

Next, we consider the case where the image is watermarked using the perceptual mask, but with a large watermark power that renders it visible. The PSNR (Peak Signal to Noise Ratio) defined as

$$PSNR = 10\log_{10}\left(\frac{\text{Maximum Possible Luminance}}{\text{Mean Squared Error}}\right) \quad (11)$$

was for this case 25 dB (note that for an invisible watermark it is customarily assumed that at least 35 dB of PSNR are necessary). The parameters for the test in this case are shown in Table 2.

| $n$ | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| $\|a_n\|$ | 0.1270 | 0.0138 | 0.0044 | 0.0004 | 0.0010 |
| $\mu_n$ | 0.0036 | 0.0021 | 0.0017 | 0.0021 | 0.0022 |
| $\sigma_n$ | 0.0041 | 0.0012 | 0.0007 | 0.0009 | 0.0008 |

**Table 2**. Fourier coefficients and noise-level parameters for the watermarked image with PSNR=25 dB.

It is clear that now $n_* = 3$, evidencing that the image was watermarked. For our final experiment, we decided to increase the PSNR to 40 dB so as to make the watermark less visible and expectedly less detectable; the results are given in Table 3.

| $n$ | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| $\|a_n\|$ | 0.0827 | 0.0081 | 0.0040 | 0.0030 | 0.0004 |
| $\mu_n$ | 0.0037 | 0.0028 | 0.0027 | 0.0029 | 0.0030 |
| $\sigma_n$ | 0.0024 | 0.0018 | 0.0017 | 0.0018 | 0.0017 |

**Table 3**. Fourier coefficients and noise-level parameters for the watermarked image with PSNR=40 dB.

We see that $n_* = 2$, still correctly suggesting that the image was watermarked. Again, we stress that these only constitute preliminary, albeit promising, results and that exhaustive testing is necessary.

## 6. CONCLUSIONS

The gradual and inevitable advance towards an all-digital world has brought about the undesirable feature of expediting the manipulation or even the fabrication of digital assets. There is then an increasing need for simple tools that allow to identify those misuses as a first step to a more detailed and costly analysis. Benford's law is an excellent candidate which, in fact, is already being used in some commercial software packages for the analysis of financial fraud. Here we have shown how a generalization of Benford's law can be employed for steganalytic purposes in images, that is, for detecting whether a certain natural image contains a hidden message. We have done so by proving for the first time that Generalized Gaussian (GG) distributions follow a generalized form of Benford's law and, furthermore, that this extends to combinations of GGs, opening the gate to video forensic applications.

Our generalization of Benford's law heavily relies on a Fourier series expansion of the data pdf in the Benford domain. This expansion had been previously used to justify convergence of an infinite product of random variables to the Benford distribution [10], but to the best of our knowledge, it

has been for the first time applied to improve the predictive value of Benford's law. Given the fact that an exponential distribution can be seen as a particular instance of a one-sided Generalized Gaussian, our results immediately extend to this case. In fact, as half-life decays of radioactive particles are known to follow an exponential distribution, it is clear that Benford's law alone is not sufficient to predict the distribution of the MSD; this might well explain the discrepancies with the Benford distribution observed by the authors in [3] and which therein had been explained by statistical variation. Other phenomena following exponential distributions, such as the arrival time between packets in Internet, can be accurately predicted by our proposed generalization.

There are several extensions to our work that are worth mentioning: on the one hand, our theoretical analysis can be adapted to include the so-called Generalized Gamma distributions, which offer slightly better image modeling capabilities. The approximation in (8) is valid for this case. Another extension concerns other domains: for instance, it has been observed that the output of the Discrete Wavelet Transform (DWT) can be also modeled by Generalized Gaussian distributions; hence it is reasonable to expect that our generalization of the Benford distribution is also applicable in this domain. Other decompositions, such as the Gabor, might also satisfy this property. At a more practical level, experimental results, not reported here, lead to conclude that our generalization also applies to speech and music signals, thus paving the way for forensic applications dealing with this kind of sources. Finally, we would like to mention that the watermark detection scheme proposed in Section 5 also finds application in deriving watermark removal attacks: in many cases, the attacker is able to gather different pieces of contents that are known to be watermarked with the same message, which is typically unknown. However, if the attacker happens to have a manageable set of message candidates, then he can try to remove a tentative message and then use our proposed test to determine whether the procedure has ended successfully.

## 7. REFERENCES


[1] F. Benford, "The law of anomalous numbers," *Proc. of the American Philosophical Society*, vol. 78, pp. 551–572, 1938.

[2] S. Newcomb, "Note on the frequency of the use of the different digits in natural numbers," *American Journal of Mathematics*, vol. 4, pp. 39–40, 1881.

[3] B. Buck, A. Merchant, and S. Perez, "An illustration of Benford's first digit law usinbg alpha decay half lives," *European Journal of Physics*, vol. 14, pp. 59–63, 1993.

[4] M. Nigrini, "A taxpayer compliance application of Benford's law," *Journal of the American Taxation Association*, vol. 1, pp. 72–91, 1996.

[5] A. Diekmann, "Not the first digit! using Benfords law to detect fraudulent scientific data," Preprint ETH, Zurich, Switzerland.

[6] A. Berger, L. Bunimovich, and T. Hill, "One-dimensional dynamical systems and Benford's law," *Transactions of the American Mathematical Society*, pp. 197–219, 2005.

[7] J.M. Jolion, "Images and Benford's law," *Journal of Mathematical Imaging and Vision*, vol. 14, no. 1, pp. 73–81, 2001.

[8] E.D. Acebo and M. Sbert, "Benford's law for natural and synthetic images," in *Proc. of the First Workshop on Computational Aesthetics in Graphics, Visualization and Imaging*, L. Neumann, M. Sbert, B. Gooch, and W. Purgathofer, Eds., Girona, Spain, May 2005, pp. 169–176.

[9] T. Hill, "Base invariance implies Benford's law," *Proc. of the American Mathematical Society*, vol. 123, pp. 887–895, 1995.

[10] J. Boyle, "An application of Fourier series to the most significant digit problem," *American Mathematical Monthly*, vol. 101, pp. 879–886, 1994.

[11] A. Hjorungnes, JM Lervik, and TA Ramstad, "Entropy coding of composite sources modeled by infinite gaussianmixture distributions," in *Proc. of the IEEE Workshop on Digital Signal Processing*, Loen, Norway, September 1996, pp. 235–238.

[12] I.S. Gradshteyn and I.M. Ryzhik, *Table of Integrals, Series, and Products, 5th Ed.*, Academic Press, San Diego, USA, 1994.

[13] C. Carslaw, "Anomalies in income numbers: Evidence of goal oriented behavior," *The Accounting Review*, vol. 63, no. 2, pp. 321–327, 1988.

[14] Ch. Schaefer, J.P. Schrapler, J.P. Muller, and G.G. Wagner, "Automatic identification of faked and fraudulent interviews in the german SOEP," *Journal of Applied Social Science Studies*, vol. 125, pp. 183–193, 2005.

[15] S. Lyu and H. Farid, "Detecting hidden messages using higher-order statistics and support vector machines," *5th International Workshop on Information Hiding*, 2002.

[16] J.R. Hernández, M. Amado, and F. Pérez-González, "DCT-domain watermarking techniques for still images: Detector performance analysis and a new structure," *IEEE Trans. on Image Processing*, vol. 9, no. 1, pp. 55–68, January 2000.