9-16-2007

BENFORD'S LAW IN IMAGE PROCESSING

Chaouki T. Abdallah

Greg L. Heileman

Fernando Perez-Gonzalez

BENFORD'S LAW IN IMAGE PROCESSING

Fernando Pérez-González †*, Greg 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 steganalysis, namely, detecting that a given image carries a hidden message.

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

1. INTRODUCTION

Benford's law of "anomalous digits" was enunciated by General Electric's physicist Frank L. Benford in 1938 [**?**], and predicts the frequency of appearance of the most significant digit (MSD) for a broad range of natural and artificial data. 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 tells that the probability that the MSD take the value $d \in \{1, 2, \cdots, 9\}$ is

$$P(d) = \log_{10}(1 + 1/d) \tag{1}$$

Since Benford's paper, many works 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 **??**. On the other hand, at a practical level, Benford's law has been shown to apply to half-life time of radioactive particles [**?**], financial data [**?**], regression coefficients, and many other types of data. Of particular interest for our purposes is the work by J.M. Jolion [**?**], 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 [**?**], 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 **??**) 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.

This work was partially funded by *Xunta de Galicia* under projects PGIDT04 TIC322013PR, PGIDT04 PXIC32202PM, and Competitive research units program Ref. 150/2006; MEC project DIPSTICK, reference TEC2004-02551/TCM, and European Commission through the IST Programme under Contract IST-2002-507932 ECRYPT.

2. BACKGROUND

In this Section we will recall some of the known properties that affect random variables in the context of Benford's law.

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

A random variable satisfying this latter property is called *strong Benford*, while the domain where Y is defined is called the *Benford domain*.

Property 2 (Scale invariance): *Suppose that X follows Benford's law; then the random variable $Z = \alpha X$ will follow Benford's law for an arbitru ary α if only if X is strong Benford.*

Besides scale invariance, it can be shown that Benford's law is also related to *base invariance*. In fact, it can be shown that a random variable is scale and base invariant if and only if it is strong Benford.

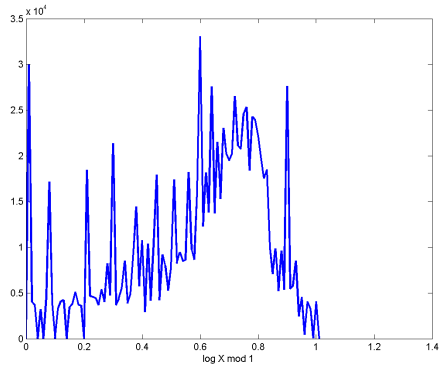
Property 3 (Product of independent random variables): *Let X be strong Benford, and let Y be another random variable indepedent of X. Then, the random variable $Z = X \cdot Y$ is strong Benford.*

The product interpretation 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. [**?**] 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 controled by an exponential distribution. Thus, if $f_X(x)$ denotes a zero-mean unit-variance Gaussian pdf, the mixture takes the form

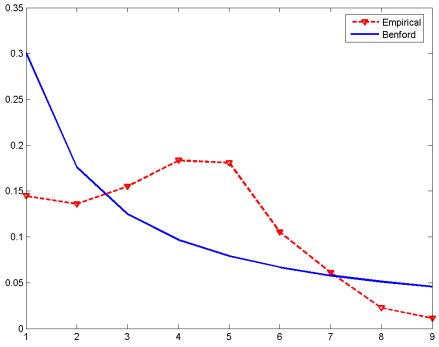
$$f_Z(z) = \int_0^\infty f_X(z|\sigma^2) f_{\sigma^2}(\sigma^2) d\sigma^2 \tag{2}$$

where $f_{\sigma^2}(\sigma^2)$ is an exponential. Interestingly, mixtures of the general form given in (**??**) can be written in such a way that Property **??** can be straightforwardly applied. Indeed, the random variable Z whose pdf is $f_Z(z)$ is obtained through (**??**) can be written as $Z = X \cdot \sqrt{\Sigma}$, with Σ 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.

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



(a)

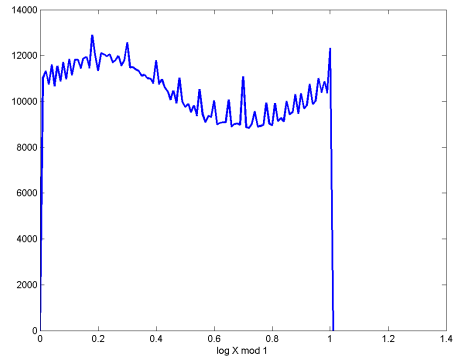


(b)

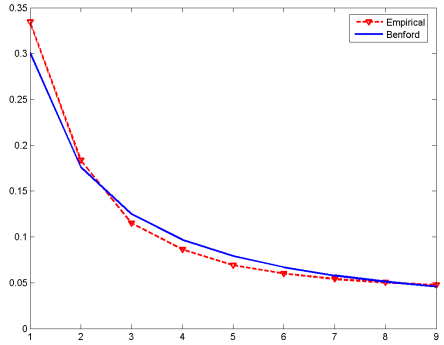
Fig. 2. Histogram of the luminance values of 'Man' in Benford ($\log_{10}$ mod 1) domain (a); Distribution of the MSD corresponding to 'Man' (b).

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 for images. Unfortunately, it is well known that image luminances 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 grey-level images do not satisfy Benford's law. To illustrate, consider the image 'Man' shown in Fig. **??** for which the histogram of the variable $\log_{10} X \mod 1$ is shown in Fig. **??**(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. **??**(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 8×8.

However, one obtains quite different results when considers the block-wise DCT transform, as it is found that the coefficients thus produced match Benford's law reasonably well. Figure **??**(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. **??**(b) represents the distribution of the MSD, which now lies much closer to Benford's distribution. The DCT block size used in these figures is 8×8; however, similar results are obtained by considering other block sizes as well as other images. A crucial observation from Fig. **??**(a) is that the histogram is not really flat, but instead can be modeled with a constant plus a sinusoidal (AC) term. This somehow suprising 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. The following fact is known about images in the DCT domain: 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 β

are expressed in terms of c and the standard deviation σ as follows

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

The parameter c is also regarded to as the *shaping factor*. Unfortunately, the parameters σ 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.

4. A FOURIER-SERIES-BASED MODEL

The sinusoidal character of the histogram in Fig. **??**(a) suggests that a Fourier representation for the pdf of the variable $\tilde{X} \triangleq \log_{10} X \mod 1$ would be plausible. We can write

$$f_{\tilde{X}}(x) = 1 + 2\sum_{k=1}^{\infty} |a_k| \cos(2\pi k x + \phi_k), x \in [0,1) \tag{4}$$

where $a_k = |a_k| e^{j\phi_k}$ is the kth Fourier coefficient. 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

$$\begin{aligned} a_n &= \int_{-\infty}^{\infty} f_{\tilde{X}}(x) e^{-j2\pi n x} dx \\ &= \frac{2A}{\beta c} e^{j\frac{2\pi n \log\beta}{\log 10}} \Gamma\left(\frac{-j2\pi n + \log 10}{c \log 10}\right) \end{aligned} \tag{5}$$

The most important observation is that the magnitude of the coefficients in (**??**) decreases very rapidly with n. In fact, it is possible to show that

$$|a_n|^2 = \prod_{k=0}^{\infty}\left(1 + \frac{(2\pi n)^2}{\log^2(10)(ck+1)^2}\right)^{-1} \tag{6}$$

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

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

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

while for a Laplacian (i.e., $c = 1$) we get the following values

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

Finally, for $c = 0.5$ we observe that 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

$$f_{\tilde{X}}(x) \approx 1 + 2|a_1| \cos(2\pi k x + \phi_k), x \in [0,1) \tag{7}$$

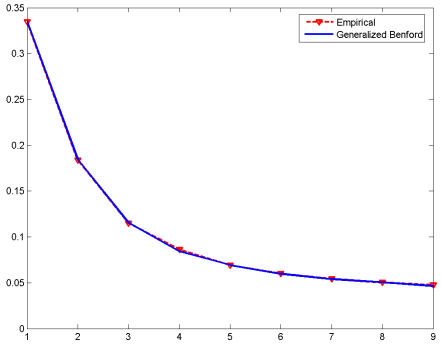


Fig. 4. Empirical digit distribution and generalized Benford's law. $a_1 = 0.067$; $\phi_1 = -1.221$ rad.

is reasonable. In Fig. **??** we plot the theoretical MSD distribution that results after using the approximation in (**??**), which leads to an excellent agreement with the empirical distribution.

So far, we have shown that a GG distribution can be closely approximated as in (**??**). However, as we have remarked, different DCT coefficients will have different parameters σ 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 Benford domain can be written as

$$\begin{aligned} f_{\tilde{X}}(x) &= 1 + \int\int \sum_{\substack{k=-\infty \\ k\neq 0}}^{\infty} a_k(c,\sigma) e^{j2\pi k x} f_{C,\Sigma}(c,\sigma) dc \cdot d\sigma \\ &= 1 + 2\mathrm{Re}\{\sum_{k=1}^{\infty} \overline{a}_k e^{j2\pi k x}\} \end{aligned} \tag{8}$$

where $\overline{a}_k$ is the mean value of a_k averaged over the joint distribution of C and Σ. 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 (**??**) is valid. Now suppose that c is such that for all i,k, $c(i,k) \leq c^+$, then for all (i,k) and all n it can be proven that $|a_n(c,\sigma)| \leq |a_n(c^+)|$, again suggesting that for values of c^+ less than 2, as is customary in practice, the approximation given in (**??**) is valid.

The previous discussion has important implications in video: if all frames of a video sequence can be modeled as in (**??**), 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.

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 [**?**], [**?**]. 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. Other recent applications in forensics include the detection of scientific data manipulation and the analysis of fabricated data in surveys.

In view of those applications, it is natural to ask whether Benford's law may find any use in image forensics. Although we foresee other specific applications, here we focus on *image steganography*. 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 is chracterized by 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 [**?**]. 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 **??** 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 μ_n and typical deviation σ_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
μ_n	0.0032	0.0029	0.0027	0.0028	0.0028
σ_n	0.0013	0.0016	0.0016	0.0017	0.0017

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

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) 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 **??**.

n	1	2	3	4	5
$\|a_n\|$	0.1270	0.0138	0.0044	0.0004	0.0010
μ_n	0.0036	0.0021	0.0017	0.0021	0.0022
σ_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 **??**.

n	1	2	3	4	5
$\|a_n\|$	0.0827	0.0081	0.0040	0.0030	0.0004
μ_n	0.0037	0.0028	0.0027	0.0029	0.0030
σ_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 consitute 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 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.

7. REFERENCES

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

[2] 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.

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

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

[5] 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.

[6] 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.

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

[8] 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.