IMAGE RECONSTRUCTION USING THE BENFORD LAW

João Sanches Jorge S. Marques

Instituto Superior Técnico / Instituto de Sistemas e Robótica
Lisbon, Portugal

ABSTRACT

It has been claimed that the first digit of real signals follows a logarithmic distribution, called Benford law. This paper shows that this distribution is a natural prior for the gradient of several types of medical images (MRI, CT, ultrasound) and proposes a reconstruction algorithm based on the Benford law which does not require any parameter tuning. Experimental results are presented to illustrate the performance of the reconstruction algorithm.

Index Terms— Biomedical imaging, Image reconstruction, Image restoration.

1. INTRODUCTION

Several methods have been proposed for denoising and reconstruction, e.g., wavelets [1], anisotropic diffusion [2] and level sets [3]. Bayesian methods [4], have been widely used to reconstruct images from noisy and distorted data, achieving remarkable results in many problems. They are based on the minimization of an energy function with two terms: a data fidelity term and a regularization term

$$E(Y,X) = -\log p(Y|X) - \log p(X) \tag{1}$$

where X is the image to be reconstructed, Y is the observed image and $p(Y|X), p(X)$ are the observation model and the prior distribution, respectively.

The choice of the prior distribution is an important issue since it influences the final result. Several priors have been proposed (e.g. using Gibbs distributions) but they can not be considered as natural priors for the class of images being considered and they often depend on hyper parameters which are difficult to estimate. Therefore, prior selection is still an open issue.

The Benford law, also known as, *First Digit Law*, was observed for the fist time in 1881 by Simon Newcomb [6]. Fifty years later the physicist Benford [7] has made exhaustive experimental tests that have confirmed the results observed by Newcomb. Basically, Benford noted that the statistical distribution of the first significant digit from a set of real life measures is not uniform, as expected, but follows a logarithmic distribution,

$$p(n) \quad = \quad k\log(1+n^{-1})$$

where k is a normalization constant. $p(n) = \{$ 0.301, 0.176, 0.125, 0.097, 0.079, 0.067, 0.058, 0.051, 0.046$\}$.

After the work of Benford several authors have presented different mathematical explanations for the Benford Law, but only in 1995, T. P. Hill [8] provided a complete mathematical formulation for the problem, including a generalization of the Benford law for the other significant digits after the first. Varian [9] has suggested using the Benford law as a test of reasonableness (naturalness) of the data.

Jolion [10] showed that the Benford law describes well the first digit of the gradient magnitude in natural images. Furthermore, he derived a prior distribution for the gradient magnitude, assuming that the first digit obeys a Benford Law, given by

$$p(g_i) = \frac{k}{g_i} \tag{2}$$

where k is a constant and g_i is the gradient magnitude at position i.

This paper shows that the first digit of the gradient magnitude in MRI, CT and ultrasound images obeys the Benford Law. Second, we propose a reconstruction algorithm based on the Benford Law which does not require the adjustment of any regularization parameter and achieves good reconstruction results.

2. PROBLEM FORMULATION

To estimate the image X, we use the *maximum a posteriori* (MAP) criterion which is equivalent to solve the following optimization problem:

$$\widehat{X} = \arg\min_X E(Y,X), \tag{3}$$

where the energy $E(Y,X)$ is given by (1).

In this paper we assume that X is a Markov random field (MRF), and therefore, the prior function, $p(X)$, is a Gibbs distribution, $p(X) = Z^{-1}e^{-\alpha U(X)}$ with

$$U(X) = \sum_i v(\delta_{iv}, \delta_{ih}), \tag{4}$$

where α is the prior parameter, Z is a normalization constant, $v(s)$ is a potential function and $\delta_{iv} = x_{i,j} - x_{i,j-1}$ and $\delta_{ih} = x_{i,j} - x_{i-1,j}$ are the first order horizontal and vertical differences on x_i, respectively.

Correspondent author: J. Sanches (jmrs@alfa.ist.utl.pt)

Assuming a Benford law and independence for the first digit of the gradient magnitudes, denoted by G, $p(G) = \prod_i p(g_i)$ where $p(g_i)$ is given by (2). Therefore,

$$p(G) = W^{-1} e^{-\sum_i \log(g_i)} \tag{5}$$

where W is a normalization constant. The gradient magnitude, g_i, at the ith pixel can be approximated by using the first order vertical and horizontal differences, $g_i = \sqrt{\delta_{iv}^2 + \delta_{ih}^2}$.

If $\alpha = 1$ and $v_i(\delta_{iv}, \delta_{iv}) = \log(g_i) \Leftrightarrow P(G) = \frac{Z}{W} P(X)$, i.e., if the first digit of the gradient magnitudes follows the Benford law, which is the case of natural images, then $p(X)$ is a Gibbs distribution with the potential function

$$v(\delta_{iv}, \delta_{ih}) = \log(g_i) = \log(\sqrt{\delta_{iv}^2 + \delta_{ih}^2}). \tag{6}$$

Table 1 compares the potential functions and influence functions associated to Benford prior, (2), with other well known priors, L_1, L_2, GGMRF [13, 12], Geman and McClure [5], Lorentzian and Hubert [11]. These functions are also displayed in Fig.1. The Benford prior is an edge preserving prior since its potential function, $v(x) = \log(|x|)$, has a sub-linear growth when x tends to infinity. Notice that, in this 2D case the potential functions are $v(\delta) = \log\sqrt{\delta^2 + k^2}$ where k is the other of the two differences involved in each clic.

In this paper, we will use $v(\delta_{iv}, \delta_{ih}) = \log(g_i + \epsilon)$ where ϵ is a constant (typically 10^{-6}) which avoids divisions by zero in homogeneous regions without any texture.

Consequently, the regularization term, by taking logarithms, is, $\log(p(X)) = -\sum_i \log(g_i + \epsilon) + C$

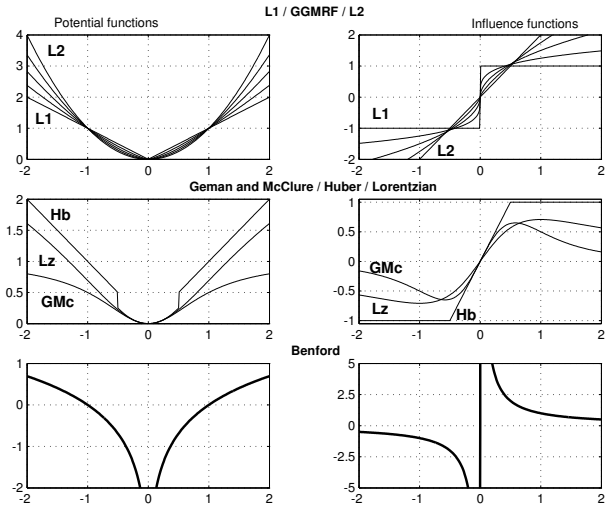

Fig. 1. Potential functions, $v(x)$, (left) and influence functions, $\psi(x) = dv(x)/dx$, (right) for different regularization priors. Generalized Gaussian MRF (GGMRF), $p = \{1, 1.25, 1.5, 1.75, 2\}$ (first line). Geman and McClure, Hubert and Lorentzian (second line). Benford prior (third line).

Let us assume for the sake of simplicity that $Y = HX + N$ where N is a Gaussian additive noise with distribution $N(0, \sigma^2 I)$. In this case, the energy is given by

	potential($v(x)$)	Influence($\psi(x)$)
L_1	$\|x\|$	$sign(x)$
GGMRF	$\|x\|^p$	$sign(x)p\|x\|^{p-1}$
L_2	x^2	$2x$
Geman	$\frac{x^2}{1+x^2}$	$\frac{2x}{(1+x^2)^2}$
Lorentzian	$\log(1 + \frac{1}{2}(\frac{x}{\sigma})^2)$	$\frac{x/\sigma}{1+\frac{1}{2}(\frac{x}{\sigma})^2}$
Hubert	$x^2/2\epsilon + \epsilon/2, \|x\| \leq \epsilon$ $\|x\|, \|x\| > \epsilon.$	$x/\epsilon, \|x\| \leq \epsilon$ $sign(x), \|x\| > \epsilon.$
Benford	$\log(\|x\| + \epsilon)$	$sign(x)/(\|x\| + \epsilon)$

Table 1. Common edge preserving priors (excluding L_2)

.

$$E(Y, X) = \frac{1}{2\sigma^2}\|HX - Y\|_2^2 + \sum_i \log(g_i + \varepsilon). \tag{7}$$

where H is a blurring operator. In this paper we will only deal with the denoising problem, so $H = I$.

3. OPTIMIZATION

The minimization of the non quadratic equation (7) is obtained by solving $\frac{dE}{dX} = 0$. In this paper we will use a bounded

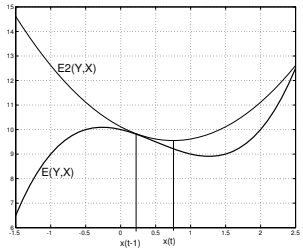

Fig. 2. Bounded optimization.

optimization algorithm proposed in [14] where the energy function is iteratively minimized. In each iteration, a quadratic energy function, E_2, is minimized such that

$$\begin{aligned} E_2(Y, X) &\geq E(Y, X) \\ E_2(Y, \hat{X}(t-1)) &= E(Y, \hat{X}(t-1)) \end{aligned}$$

where $\hat{X}(t-1)$ is the estimate of X at iteration $t-1$ (see Fig.2). The minimization of E_2 leads to a new estimate of X. This procedure is iteratively performed until convergence is achieved.

As shown in [14] the equation (7) can be minimized by iteratively solving the following equation

$$E_2(Y, X) = \frac{1}{2}\|X - Y\|^2 + \sigma^2 \Delta^T D \Delta \tag{8}$$

where $\Delta = \{\delta_i\}$ is a column vector containing all vertical and horizontal differences $x_p - x_q$ for $\{p, q\} \in N$, where N denotes the set of adjacent nodes and D is a diagonal matrix.

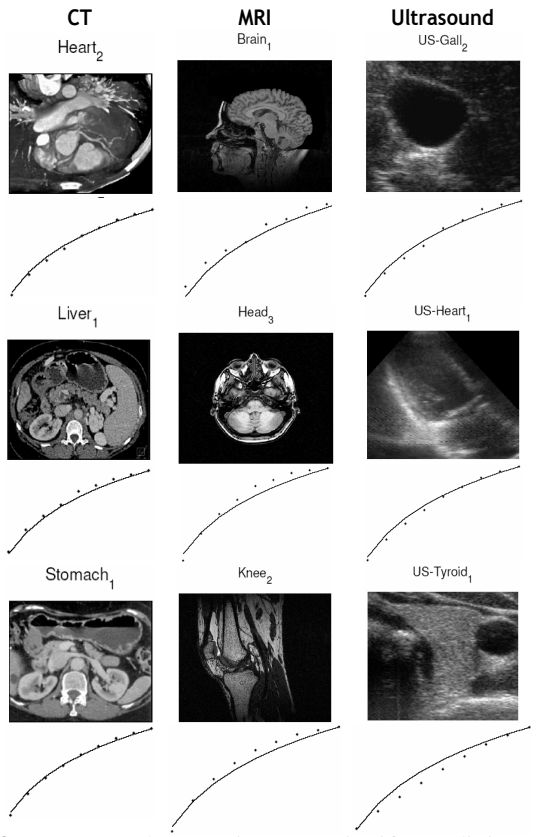

Fig. 3. Images and respective most significant digit cumulative distributions of the gradient magnitude (experimental histogram and Benford law), for CT (first column), MRI (second column) and US (third column).

The difference vector Δ can be computed as $\Delta = \Theta X$ where Θ is an appropriate matrix. The elements of diagonal matrix D are given by

$$d_i = -\frac{1}{\delta_i p(\delta_i)} \frac{dp(\delta_i)}{d\delta_i} \tag{9}$$

where $p(\delta_i) = \frac{K}{g(\delta_i)+\epsilon}$, given by (2), thus,

$$\frac{dp(\delta_i)}{d\delta_i} = -\frac{K\delta_i}{g(\delta_i)(g(\delta_i)+\epsilon)^2} \Rightarrow \tag{10}$$

$$d(\delta_i) = \frac{1}{g(\delta_i)(g(\delta_i)+\epsilon)} \approx \frac{1}{g^2(\delta_i)+\epsilon} \tag{11}$$

The minimum of (8) is given by

$$X_{t+1}^T = (I + \sigma^2 \Theta^T D \Theta)^{-1} Y \tag{12}$$

4. EXPERIMENTAL RESULTS

In this section we present two types of experiments. First we test if the Benford Law is valid in a data set of 400 real medical images, using the Kolmogorov-Smirnov conformity test. In the second set of experiments we apply the denoising algorithm to several medical images, corrupted by additive Gaussian noise.

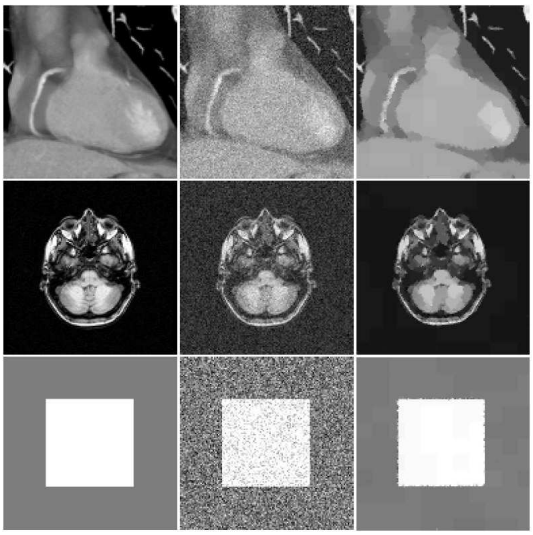

Fig. 4. Denoising results with artificially corrupted images. CT(first line), MRI(second line) and Synthetic(third line) images. (first column) Clean original images, (second column) artificially corrupted with gaussian additive noise ($\sigma = 20$) and (last column) denoised images.

4.1. Benford law conformity tests

This section applies conformity tests to the first digit of the gradient magnitude for two different hypotheses: Benford law and uniform distribution. To measure the degree of fitness between the theoretical and experimental curves we use the level of confidence on the null hypotesis, H_0, given by the Kolmogorov-Smirnov statistical test. We have also decided to compare the experimental data with the uniform distribution because this one is more intuitive for the most significant digit (despite this intuitiveness is wrong as proven in [10]). Besides, the comparison with the uniform law help us to evaluate the degree of closeness with the Benford law.

$P_e = 1 - P_{H_0}$ is the probability of rejection of the null hypothesis, H_0, which is, the hypothesis of the data have been generated by the Benford/Uniform law. For the Kolmogorov-Smirnov test $P_{H_0} = Q_{KS}(\lambda)$ [1] where $\lambda = (\sqrt{N} + 0.12 + 0.11/\sqrt{N})D$ and $N = 9$ is the number of data points. $D = max|p(n) - h(n)|$ is the maximum absolute difference between the theoretical distribution, $p(n)$ and the experimental histogram, $h(n)$, for $1 \leq n \leq 9$. Table 2 shows the geometric mean of P_e for the most significant digit law, computed over three different sets of medical images, CT (60 images), MRI (409 images) and US (7 images). The resulting experimental histograms are compared with the theoretical Benford (second column of table 2) and Uniform (third column of table 2) laws. Figure 3 displays examples of medical images and the

[1] $Q_{KS}(\lambda) = 2\sum_{j=1}^{\infty}(-1)^{j-1}e^{-2j^2\lambda^2}$

P_e	Benford	Uniform
CT	$7.06E-14$	$4.41E-01$
MRI	$1.06E-11$	$6.84E-01$
US	$2.38E-13$	$2.34E-01$

Table 2. Kolmogorov-Smirnov tests: geometric mean of the error probability, P_e, computed over 476 images.

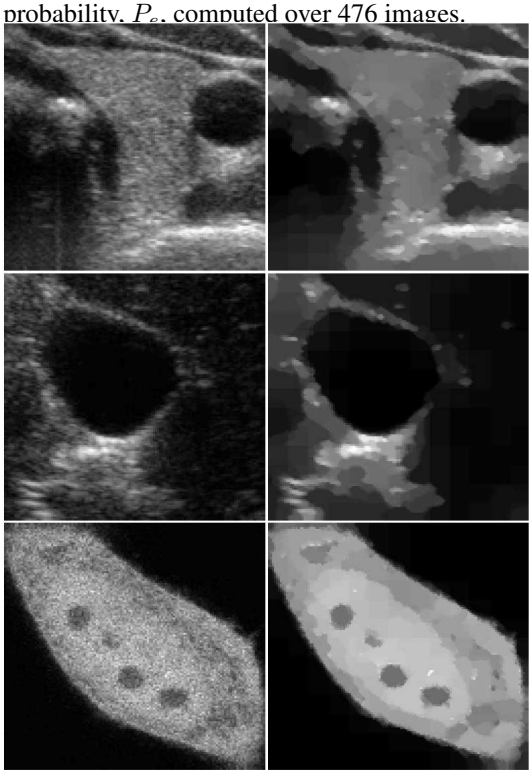

Fig. 5. Denoising results using real noisy images. (first and second columns) ultrasound images and (third column) fluorescent confocal microscopy image os a cell.

correspondent cumulative distributions of the most significative digit, as well, the theoretical Benford law. These images, show a good fit of the experimental data with the Benford law. This fitness is reenforced with the Kolmogorov-Smirnov P_e measures presented in the table 2, which shows a better adjustment to the Benford law than the uniform distribution.

4.2. Denoising

Here, two types of experimental results are presented. First, three low noise images are artificially corrupted with additive Gaussian noise. These noisy images are denoised using the algorithm (12). Fig. 4 displays clean CT, MRI and synthetic images (first column), the correspondent corrupted images (second column) and the denoised images in the last column. The signal to noise ratio, SNR, is $22.6dB$ for the CT image, $16.4dB$ for the MRI image and $29.8dB$ for the synthetic case.

In the second type of experiments the reconstruction algorithm is applied to three natural noisy images. Two ultrasound images and one fluorescent image from confocal microscopy are used due its high level of noise. Fig. 5 shows the original images (first column) and the denoised images (second column).

5. CONCLUSIONS

This paper discusses the application of the Benford Law to the representation of medical images: CT, MRI, US. It is shown that the Benford Law provides a valid distribution for the most significant digit of the gradient magnitude, extending the results in [10]. This allows to derive a natural prior without hyper parameter to describe the distribution of the image gradient, in medical images.

A denoising algorithm based on the Benford prior is developed and tested with medical images leading to good image estimates. These results were obtained without using any hyper parameter which is difficult to estimate in real operations.

6. REFERENCES

[1] Chang, G., Yu, B., and Vetterli, M. Adaptive wavelet thresholding for image denoising and compression, IEEE Trans. Image Processing, 9 (2000), 1532-1546.

[2] C. Riddell, H. Benali, I. Buvat, Diffusion Regularization for Iterative Reconstruction in Emission Tomography, IEEE Trans. Nuclear Science, VOL. 52, 669-675, 2005

[3] T. Deschamps, R. Malladi, I. Ravve, Fast Evolution of Image Manifolds and Application to Filtering and Segmentation in 3D Medical Images, IEEE Trans. Visualization and Computer Graphics, Vol. 10, 525-535, 2004

[4] M. R. Banham, A. K. Katsaggelos, Digital Image Restoration, IEEE Signal Processing Magazine, Vol. 14, n 2, March 1997.

[5] S. Geman, D. McClure, Bayesian Image Analysis: An Application to single photon emission tomography, Proc. Statist. Comput. Sect., Amer.Stat.Assoc., pp.12-18, 1985.

[6] Newcomb S., Note on the frequency of use of different digits in natural numbers, Amer. Math. vol.4, pp.39-40, 1881.

[7] Benford F., The law of anomalous numbers, Proc. Amer. Phil. Soc. vol.78, pp.551-572, 1938.

[8] Hill T.P., The Significant-Digit Phenomenon, American Matehematical Monthly vol.102, pp.322-327, 1995.

[9] Varian H., Benford's Law, Amer. Stat. vol.26, pp.65-66, 1972.

[10] Jolion J.M., Images and Benford's Law, Journal of Math. Imaging and Vision, vol.14, pp.73-81, 2001.

[11] Black M. J., A. Rangarajan, On the unification of line processes, outlier rejection and robust statistics with applications in early vision, Journal of Comp. Vis, 1996.

[12] C.R.Vogel, M.E.Oman, Fast, Robust Total Variation-Based Reconstruction of Noisy, Blurred Imges, IEEE Trans. on Image Processing, Vol.7, no.6, June 1998.

[13] Bouman C., Sauer K., A Generalized Gaussian Image Model for Edge-Preserving MAP Estimation, IEEE Trans. Image Process., vol. 2, no. 3, pp. 296-310, July 1993.

[14] J.M.Bioucas-Dias, Fast GEM wavelet-based image deconvolution algorithm, IEEE ICIPC03, vol.2,pp.961-964,2003.