ON THE STATISTICS OF NATURAL STOCHASTIC TEXTURES AND THEIR APPLICATION IN IMAGE PROCESSING

Ido Zachevsky and Yehoshua Y. Zeevi

Technion, Israel Institute of Technology
Haifa, 32000, Israel

ido@tx.technion.ac.il, zeevi@ee.technion.ac.il

ABSTRACT

Statistics of natural images has become an important subject of research in recent years. The highly kurtotic, non-Gaussian, statistics known to be characteristic of many natural images are exploited in various image processing tasks. In this paper, we focus on natural stochastic textures (NST) and substantiate our finding that NST have Gaussian statistics. Using the well-known statistical self-similarity property of natural images, exhibited even more profoundly in NST, we exploit a Gaussian self-similar process known as the fractional Brownian motion, to derive a fBm-PDE-based single-image superresolution scheme for textured images. Using the same process as a prior, we also apply it in denoising of NST.

***Index Terms*—** Image texture enhancement, superresolution, denoising, self-similarity, fractional Brownian motion, natural image statistics.

1. INTRODUCTION

Statistics of natural images have been the subject of intensive studies in recent years [1–3]. With the increased use of statistical image enhancement algorithms, suitable priors play a crucial role in the enhancement and restoration of images, especially in ill-posed problems where severely degraded images are at hand.

Various studies have consistently shown that natural images exhibit non-Gaussian behaviour. This has been observed by inspecting the 1D, 2D, or joint histogram of the wavelet coefficients of an image [4, 5]. These histograms, evaluated on numerous images, provide a good indication of the statistical nature of the images. Rather than Gaussianity, natural images exhibit highly kurtotic behaviour, indicated by heavy tails in both 1D and joint distributions [6]. Many models capture this behaviour successfully, such as Gaussian scale mixtures (GSM) [4, 5] or generalized normal [7].

While previous studies attempt to capture the entire range of natural images, we consider natural stochastic textures (NST), that are abundant in natural images [8]. Many natural textures, such as sand, gravel, grass, grove and others exhibit fine details that are severely degraded by cameras' PSFs, sampling and noise. Such textures are not well represented by current models, and therefore enhancement algorithms, such as superresolution or denoising, do not perform well on NST. This problem is intensified when L_2- or L_1-based methods are implemented, since these reward smooth or piecewise-smooth images. This is addressed in recent studies attempting to figure out whether images are of bounded variation space [9].

Contrary to recent findings indicating that natural images are in general non-Gaussian, NST exhibit, in fact, Gaussian statistics. Coupled with the self-similarity property, a model can be proposed, based on the fractional Brownian motion (fBm) process, which is the only Gaussian self-similar process (in 1D). In this study we encorporate this model in a PDE framework to obtain a novel single-image superresolution (SR) scheme, and use the fBm as a prior for a denoising scheme.

2. STATISTICS OF NATURAL STOCHASTIC TEXTURES

NST are an important part of a natural image. We briefly outline the methods of studying the statistics of natural images, and provide results that substantiate our observation that NST are Gaussian, in contrast with the known statistics of general natural images [4, 6]. The test images were taken from the VisTex texture database [10].

The images were analyzed in the wavelet domain, using steerable pyramid wavelets, thus allowing for different orientations as well as different scales (see [4]; other wavelets yield similar results). Marginal histograms of wavelet coefficients were extracted for different scales, as well as joint histograms of pairs of pixels from adjacent scales or orientations.

In [6] and elsewhere, natural images were found to be of highly distributed kurtotic values. A known distribution

This research program has been supported by the Technion Ollendorff Minerva Center for Vision and Image Sciences, and by a special grant awarded to Y.Y.Z by the Minerva Foundation. Ido Zachevsky has been an Ollendorff Fellow. The authors thank Prof. Armin Schwartzman, NSCU.

978-1-4799-2893-4/14/$31.00 ©2014 IEEE

which fits 1D empirical distribution of images is the generalized normal distribution (or generalized Laplace distribution), which has the following pdf:

$$p(x) = c(\alpha, \beta) \exp(-(|x - \mu|/\alpha)^{\beta}), \tag{1}$$

with the normalizing constant, $c(\alpha, \beta) = \frac{\beta}{2\alpha\Gamma(1/\beta)}$. The μ and α parameters correspond to the mean and variance respectively, where the variance is defined explicitly by α. The β parameter determines the kurtosis of the distribution, whose excess kurtosis is defined by:

$$K = \frac{\Gamma(5/\beta)\Gamma(1/\beta)}{\Gamma(3/\beta)^2} - 3. \tag{2}$$

$\beta = 2$ corresponds to the normal distribution with zero excess kurtosis. The range $0 < \beta < 2$ corresponds to distributions with high kurtosis. Eq. (2) shows that the excess kurtosis is defined nonlinearly with β; values of $\beta > 1$ correspond to much smaller kurtosis relative to $\beta < 1$.

The generalized normal distribution has been successfully used as a distribution for wavelet coefficients of empirical distributions of natural images for $\beta \in [0.5, 1]$ [6, 7], which correspond to leptokurtic behaviour. We propose the normal distribution model for the class of stochastic textures, obtained for $\beta = 2$ in Eq. (1). Out of the numerous methods that determine the more suitable distribution, we use the Kullback-Leibler (KL) divergence, $D_{KL}(f_1 \| f_2) = \int_{\mathbb{R}} f_1(x) \log\left(\frac{f_1(x)}{f_2(x)}\right) dx$, where $f_1(x)$ is the empirical density and $f_2(x)$ is the density according to the evaluated model [11].

We first assume the image dataset fits a leptokurtic distribution, the generalized normal, and estimate its β value according to the mean of the sample kurtosis. The result is $\hat{\beta} = 0.711$. We then propose two distributions: $p_n(x)$, a normal distribution, and $p_g(x)$, a generalized normal, with $\beta = \hat{\beta}$. Equipped with the two distributions as possible models, the KL divergences for the empirical distribution and each of the models were calculated. Out of 1914 test images, 620 (32%) had lower KL divergence for the normal distribution, $p_n(x)$, indicating that the normal distribution better describes the data. The test images were obtained by dividing all images (of size greater than 256×256) in the VisTex database to 256×256 sized images.

A second test was performed, with individual values of β, each estimated from an individual image. Inspecting all values for this parameter, 19% were above a threshold kurtosis value, which was chosen as the average between the Gaussian distribution kurtosis and the Laplace distribution kurtosis. The latter has the lowest kurtosis for known image models, with $\beta = 1$ and excess kurtosis of 3. In this case as well, we observe a significant number of images described better by the Gaussian distribution. Three individual images, along with their statistics, are displayed in Fig. 1, demonstrating the different statistics of NST in contrast with the statistics of general natural images.

2.1. Self similarity

The self-similarity property is an important property of natural images. For Gaussian processes, it can be evaluated by performing log regression on the variance of the increments in the image domain. This is due to the fact that the only Gaussian process exhibiting self-similarity is the fractional Brownian motion, whose covariance function is known [12]. The log of the variance of the increments of order τ is given by:

$$y_H(\tau) = 2Hx(\tau) + b_H, \tag{3}$$

where $y_H(\tau) = \log \sigma^2(H, \tau)$ is the measurement for each τ, $x(\tau) = \log \tau$, and $b_H = \log \sigma_B^2(H)$ is a known parameter. H is the Hurst parameter, controlling the self-similarity of the process. Performing this regression on the images found to be Gaussian in the above analysis yielded an R^2 value with median of 0.96, which indicates a realiable result and reassures that NST are Gaussian and self-similar. The median was chosen to minimize effects of outliers.

This result is important for image enhacement, as NST often exhibit fine details which are easily corrupted by blurring, decimation and noise in practical image acquisition processes. It is therefore imperative to have an accurate model, and treat NST in a separate domain than other types of textures or cartoon images.

3. SUPERRESOLUTION

Consider the following superresolution problem: A high-resolution (HR) image is degraded by a blurring filter, and it is subsequently subsampled to create the available low-resolution (LR) image. The inverse process, of obtaining an HR image is known as superresolution. Let $X(\eta_1, \eta_2)$ and $Y(\eta_1, \eta_2)$ denote the original (HR) image and observed (LR) image, respectively. The imaging model can be represented as follows:

$$Y(\eta_1, \eta_2) = \mathcal{D}\left((Y * b)(\eta_1, \eta_2)\right), \tag{4}$$

where $\mathcal{D}$ is the subsampling operator and $b(\eta_1, \eta_2)$ is a non-invertible blur kernel of limited spatial support. The decimation operator introduces aliasing and renders the SR problem to become severely ill-posed.

Unlike the classical multi-frame SR problem [13], in single-image SR only a single measurement is available. It can be formally stated as follows:

$$\hat{X}(\eta_1, \eta_2) = \arg\min_{X \in \mathcal{X}} \|Y(\eta_1, \eta_2) - \mathcal{D}\left((X * b)(\eta_1, \eta_2)\right)\|_p. \tag{5}$$

The SR image, $\hat{X}(\eta_1, \eta_2)$, thus obtained is the best one in that it yields the smallest error relative to the original image (ground truth) for a desired metric [14, 15].

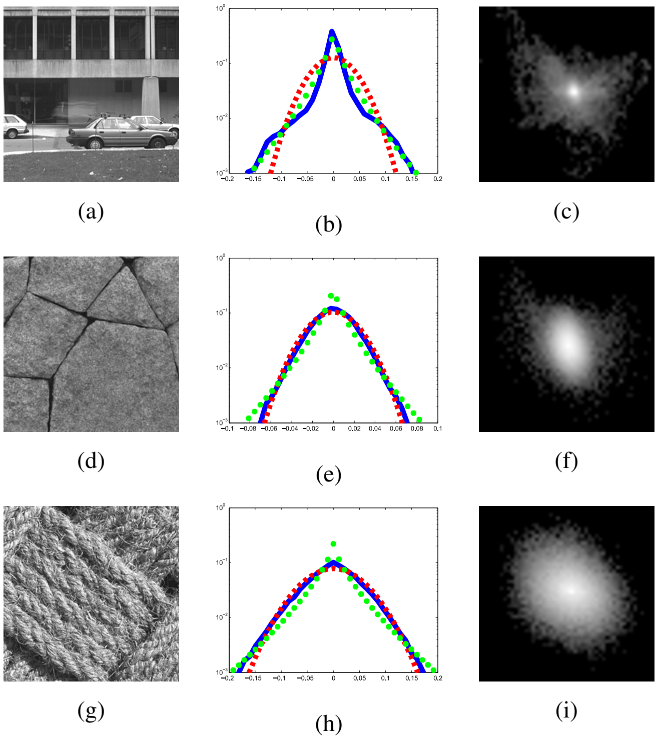

Fig. 1: Examples of general image and NST statistics: (a) A natural image. (b) Its marginal wavelet coefficient distribution, with the Gaussian model (red-dashed) and the leptokurtic model (green-dotted). (c) Its joint (2D) distribution. The KL divergences for the Gaussian and leptokurtic models were 0.58 and 0.08 respectively, indicating non-Gaussianity. (d) NST: The KL divergences were 0.03 and 0.11 for the Gaussian and the leptokurtic models respectively. (g) Another NST: The KL divergences were 0.02 and 0.15 respectively. For NST, the Gaussian model describes the images better. (e) and (h) show the empirical, Gaussian and leptokurtic distribution fits (in blue, red-dashed and green-dotted respectively). The Gaussianity of the NST is also apparent in the 2D histogram shapes in (f) and (i), compared with the non-Gaussian one, (c).

3.1. A model for natural stochastic textures

Based on the Gaussianity and self-similarity of NST, we propose a texture model based on the fBm (the only process which is both Gaussian and self-similar), and on the phase of the degraded image:

$$X = X_0 + V, \tag{6}$$

where

$$\begin{aligned} X_0 &= \alpha(Y * H_{LP}) + (1-\alpha)(W * H_{HP}), \\ W &= \mathcal{F}^{-1}\{|\mathcal{F}\{U\}| \exp(j\angle\mathcal{F}\{Y\})\}, \end{aligned} \tag{7}$$

U is a 2D fBm realization with a suitable parameter H, Y is the degraded image, V is the model noise and $\alpha \in (0,1)$. H_{LP} and H_{HP} are lowpass and highpass filters respectively. The parameter $H \in (0,1)$ controls the roughness of the fBm image. The image X_0 has three main properties: First, X_0 and the degraded image, Y, have the same low frequencies. Second, the high frequency part of X_0 has the same phase as the degraded image, derived from the well-established importance of phase in natural images [16,17]. The third property is the magnitude of the high frequency part of X_0, derived from a realization of an fBm with $H = \hat{H}$, where $\hat{H}$ is an estimation of the original H (derived from the degraded image). Its purpose is to exploit a realization with fBm high-frequency details for reconstruction of missing frequency magnitudes in the degraded image [14].

3.2. fBm-PDE-based superresolution scheme

A naive optimization scheme, based on the proposed model, is as follows:

$$\hat{X} = \arg\min_X \alpha\|Y - DBX\|^2 + \beta\|X_0 - X\|^2, \tag{8}$$

indicating that the optimal solution yields the closest to the proposed NST model and the image degradation model. However, due to the fine details exhibited in NST, we use a PDE-regularized scheme, adapted for texture enhancement as well.

The anisotropic diffusion can be used as a regularizer for image optimization problems, such as deblurring or superresolution. In this case, the cost function is:

$$\mathcal{L} = \int_\Omega \alpha(Y - DBX)^2 + \beta(X_0 - X)^2 + \gamma\Psi(\nabla X + \delta\nabla Y_\phi)dxdy, \tag{9}$$

where the first two terms are the reaction terms and the last term is the diffusion regularization term. This is an extension of the diffusion-based deconvolution [18]. Applying gradient descent on the Euler-Lagrange equation of this functional yields the diffusion flow. The diffusion flow, along with the modified reaction term, recovers degraded information, to yield an SR scheme for NST. Y_ϕ, or the empirical image, is a novel term, derived from the statistical structure of the degraded image. The first and second order autocorrelation of the increments of the degraded image increments are used to yield a structure function [19], from which a random field with stationary increments, suitable for the internal structure of the image, is generated. Although the structure function is built from the statistics of the degraded image, under the scale invariance (self-similarity) assumption, these represent the high-quality image as well.

As the diffusion flow preserves structure in the image, it also inherently smoothes low gradients. Introducing the image Y_ϕ to the diffusion tensor assists in preserving the recovered fine details and prevents their undesired smoothing. The stopping condition for the diffusion is determined by the H parameter; using the estimated original $\hat{H}$, the final image is obtained when H for the current iteration is the closest to $\hat{H}$.

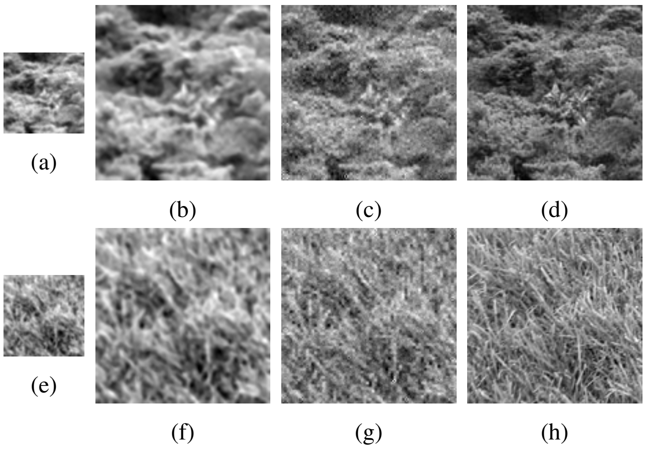

Fig. 2: Two examples of SR. (a and e): LR image. (b and f): Self-similarity-based SR, 15.14dB and 18.98dB PSNR respectively. (c and g): fBm-PDE-based SR (proposed algorithm), 19.49dB and 20.15dB PSNR respectively. (d and h) the original image (ground truth).

3.3. Results

The proposed SR algorithm was evaluated on NST images and compared with a state-of-the-art, self-similarity-based single-image SR algorithm [20]. In all the test cases such as the two shown Fig. 2, the textures were recovered by the proposed algorithm. We noticed lack of improvement by the application of existing methods. This is due to the low-gradient structure of textures, where contour enhancement is not sufficient for successful texture restoration. Furthermore, contour emphasis is undesirable in such images, contrary to the common demand in enhancement of general images, on which the same algorithm yields much better results.

Each image was blurred by a Gaussian kernel with $\sigma = 1.5$ and decimated by a ratio of 2 in both axes. The PSNR values are also provided for comparison (Fig. 2).

4. DENOISING

Denoising of NST is in particular challenging due to the fine structure of the textures. The fBm prior can be encorporated into any prior-based method. As a first attempt we chose to use optimal MMSE estimation, which yields a convenient and well-known linear estimator, due to the fact that both the image and the noise are Gaussian.

Thirty images were chosen arbitrarily out of the set of images shown to be Gaussian and 64×64 patches were extracted for the purpose of the test. Their H parameter was estimated and the reference image was obtained.

Each image was blurred by a small Gaussian filter with a $\sigma = 0.5$ and noise was added so that the blurred-PSNR varied from 10dB to 20dB. The recovered PSNR was averaged for all images. Despite using a naive scheme, we see that in terms

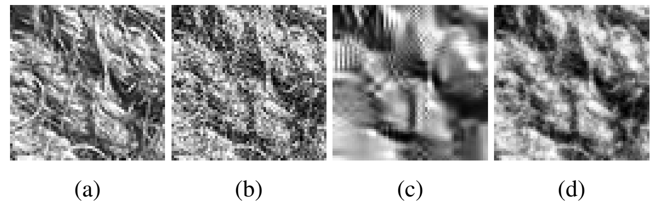

Fig. 3: Example of denoising. (a): The original image. (b) Noisy image, obtained from (a) by blurring with a Gaussian blur with $\sigma = 0.5$ and contaminating with white Gaussian noise so that the blurred PSNR is 11dB. (c): BM3D denoised image, 14.73dB PSNR. (d): MMSE fBm-based denoised image, 17.60dB PSNR.

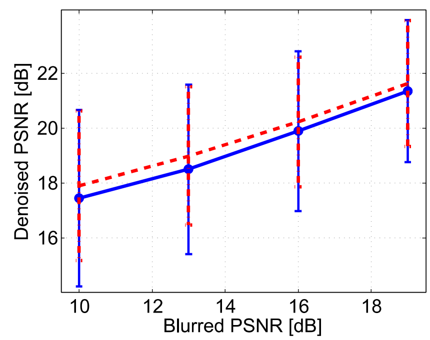

Fig. 4: Denoising results. The MMSE estimate with fBm prior (red-dashed) obtained similar results to the BM3D denoising algorithm (blue) in terms of PSNR, when performed on NST. Each PSNR value is the mean for a test of 30 images and the error bars indicate the standard deviation.

of both PSNR values and visual inspection (Fig. 3; for more examples see http://tx.technion.ac.il/~ido), the fBm-based denoised images yielded similar results to those of BM3D, a state-of-the-art denoising algorithm (Fig. 4) [21]. This is due to the undelying model, which recovered fine details rather than smoothing artifacts and preserving edges. It should be noted that the fBm-based denoising is by far more computationally efficient, as it is a linear estimator.

5. CONCLUSIONS

NSTs, unlike natural images in general, exhibit Gaussian statistics. Coupled with self-similarity, the fBm can be exploited as a prior model in various image processing problems. An fBm-based SR scheme shows satisfactory results, and even a naive MMSE linear estimator-based denoising already yields results similar to those obtained by means of a recent and more complex denoising algorithm [21]. Other fBm-based image processing schemes, as well as additional applications such as image matching and tampering detection, are currently under investigation.

6. REFERENCES

[1] A. B. Lee, D. Mumford, and J. Huang, "Occlusion Models for Natural Images: A Statistical Study of a Scale-Invariant Dead Leaves Model," *IJCV*, vol. 41, pp. 35–59, 2001.

[2] P. Moulin and J. Liu, "Analysis of Multiresolution Image Denoising Schemes Using Generalized Gaussian and Complexity Priors," *Inf. Theory, IEEE Trans.*, vol. 45, no. 3, pp. 909–919, 1999.

[3] D. Mumford and B. Gidas, "Stochastic models for generic images," *Q. Appl. Math.*, vol. LIX, no. 1, pp. 85–111, 2001.

[4] M. J. Wainwright and E. P. Simoncelli, "Scale Mixtures of Gaussians and the Statistics of Natural Images," in *Adv. Neural Inf. Process. Syst.*, vol. 12. MIT Press, May 1999, pp. 855–861.

[5] J. Portilla, V. Strela, M. J. Wainwright, and E. P. Simoncelli, "Image denoising using scale mixtures of Gaussians in the wavelet domain," *Image Process. IEEE Trans.*, vol. 12, no. 11, pp. 1338–1351, Jan. 2003.

[6] A. Srivastava, A. B. Lee, and E. P. Simoncelli, "On Advances in Statistical Modeling of Natural Images," *J. Math. Imaging Vis.*, vol. 18, pp. 17–33, 2003.

[7] E. P. Simoncelli and E. H. Adelson, "Noise removal via Bayesian wavelet coring," in *Image Process. Int. Conf.*, 1996, pp. 379–382.

[8] W.-C. Lin, J. Hays, C. Wu, V. Kwatra, and Y. Liu, "A comparison study of four texture synthesis algorithms on regular and near-regular textures," Carnegie Mellon University, Tech. Rep., 2004.

[9] Y. Gousseau and J.-M. Morel, "Are Natural Images of Bounded Variation?" *SIAM J. Math. Anal.*, vol. 33, no. 3, pp. 634–648, Jan. 2001.

[10] R. Pickard, C. Graszyk, S. Mann, J. Wachman, L. Pickard, and L. Campbell, "Vistex database," *Media Lab., MIT, Cambridge, Massachusetts*, 1995.

[11] S. Kullback, *Information theory and statistics*. Courier Dover Publications, 1997.

[12] B. Mandelbrot and J. Van Ness, "Fractional brownian motions, fractional noises and applications," *SIAM review*, vol. 10, no. 4, pp. 422–437, 1968.

[13] M. Elad and A. Feuer, "Restoration of a single superresolution image from several blurred, noisy, and undersampled measured images," *IEEE Trans. Image Process.*, vol. 6, no. 12, pp. 1646–1658, Jan. 1997.

[14] I. Zachevsky and Y. Y. Zeevi, "Superresolution of self-similar textures," *CCIT Rep. 838. EE Pub, Tech. Isr. Inst. Technol.*, vol. 1795, 2013.

[15] ——, "Single-image superresolution of self-similar textures," in *IEEE Int. Conf. Image Process.*, Sep. 2013, pp. 952–956.

[16] A. V. Oppenheim and J. S. Lim, "Importance of Phase in Signals," *Proc. IEEE*, vol. 69, no. 5, pp. 529–541, 1981.

[17] J. Behar, M. Porat, and Y. Y. Zeevi, "Image Reconstruction from Localized Phase," *IEEE Trans. Signal Process.*, vol. 40, no. 4, pp. 736–743, 1992.

[18] M. Welk, D. Theis, T. Brox, and J. Weickert, "PDE-based deconvolution with forward-backward diffusivities and diffusion tensors," *Scale Sp. PDE Methods Comput. Vis.*, pp. 585–597, 2005.

[19] B. Pesquet-Popescu and P. Larzabal, "Synthesis of nonstationary fields with stationary increments," in *Image Process. Its Appl. Sixth Int. Conf.*, vol. 1, Jul. 1997, pp. 303–307.

[20] D. Glasner, S. Bagon, and M. Irani, "Super-resolution from a single image," in *2009 IEEE 12th Int. Conf. Comput. Vis.*, Sep. 2009, pp. 349–356.

[21] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising with block-matching and 3D filtering," *Proc. SPIE*, vol. 6064, pp. 606 414–606 414–12, Feb. 2006.