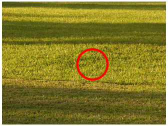
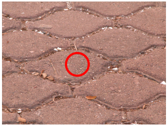
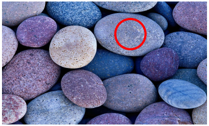
Model-based Color Natural Stochastic Textures Processing and Classification

Ido Zachevsky and Yehoshua Y. Zeevi

GlobalSIP 2015

What are Natural Stochastic Textures (NST)?

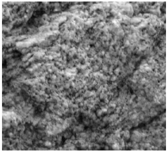

- Segments of natural images are rich in details.
- Considered to be realizations of random processes.

Fractional Brownian motion (fBm) model [Mandelbrot and Van Ness, 1968, Zachevsky and Zeevi, 2014]:

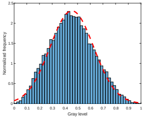
- A self-similar, isotropic, random process with stationary increments [Mandelbrot and Van Ness, 1968, Pesquet-Popescu and Vehel, 2002].
- Covariance: $R(t,s) = \sigma_H^2 \left(|t|^{2H} + |s|^{2H} - |t-s|^{2H} \right)$.
- $H \in (0,1)$: defines the regularity of the process.
- Used successfully for texture enhancement [Zachevsky and Zeevi, 2014].

NST color model (I)

- The channels of color NST are correlated.
- Higher correlation in small patches, better fractal behavior in larger patches.

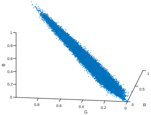


Grayscale NST



Gaussian distribution

Color NST



Correlated channels

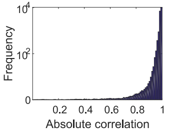
⋆ $\rho(R, G) = 0.94$, $\rho(R, B) = 0.84$.

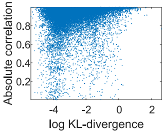
NST Color model (II)

- Patch-based analysis.
- Each RGB patch is a linear combination of a prototype fBm patch, P_0:

$$P_i(\eta_1, \eta_2) = a_i P_0(\eta_1, \eta_2) + b_i.$$

- Each patch is described by the parameters $\Theta_k = \left(\{a_i^k, b_i^k\}_{i=1}^3, H^k\right)$.





- Analysis performed on images from the McGill texture database [Olmos and Kingdom, 2004] and the VisTex database [Pickard et al., 1995].
- Patch size used: 32×32.
- High absolute correlation is indicated between patch color channels.

Deconvolution scheme

- $\mathcal{Q} = \{Q_i(\eta_1, \eta_2)\}_i$: a blurred-noisy image patch.
- Degradation model:

$$Q_i(\eta_1, \eta_2) = (P_i * B)(\eta_1, \eta_2) + N_i(\eta_1, \eta_2), i \in \{1, 2, 3\},$$

- B is a blur filter, and each $N_i(\eta_1, \eta_2) \sim \mathcal{N}(0, \sigma_N^2)$ is an independent noise image.
- Deviations from the linear model are allowed by adding model noise:

$$P_i(\eta_1, \eta_2) = a_i P_0(\eta_1, \eta_2) + (b_i + \epsilon_{b,i}) + \epsilon_i(\eta_1, \eta_2),$$

- $\{\epsilon_{b,i}\}_i$ are independent normal variables with variance $\sigma^2_{\epsilon,b,i}$, and $\{\epsilon_i(\eta_1, \eta_2)\}_i$ is an i.i.d normally distributed image with variance $\sigma^2_{\epsilon,i}$. These items allow deviations from the model by introducing model noise terms.

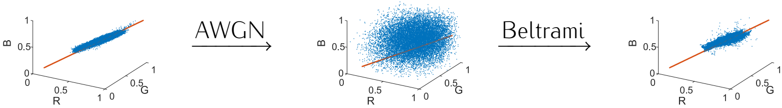
Two-stage deconvolution:

1. Parameter estimation, $\Theta = (\{a_i, b_i\}_i, H)$.
2. Maximum-a-posteriori (MAP) estimation.

Estimation (I)

Parameter estimation

- The variance of the image increments for known (α, β), $\mathrm{var}(\boldsymbol{Q}(\eta_1, \eta_2) - \boldsymbol{Q}(\xi_1, \xi_2)) = \alpha ||(\eta_1, \eta_2) - (\xi_1, \xi_2)||^{2H} + \beta$, is estimated by regression of the increments' sample variance, thus obtaining $\boldsymbol{H}$ as the power exponent.
- The color model parameters, $(\boldsymbol{a}_i, \boldsymbol{b}_i)$, defining a patch, $\boldsymbol{P}_i(\eta_1, \eta_2) = \boldsymbol{a}_i \boldsymbol{P}_0(\eta_1, \eta_2) + \boldsymbol{b}_i$, are estimated using Beltrami flow [Sochen et al., 1998].
- The Beltrami flow correlates RGB image gradients by using a diffusion flow.



- Estimation of $(\boldsymbol{a}_i, \boldsymbol{b}_i)$ is performed on the Beltrami flow processed image.
- Optimal $(\boldsymbol{a}_i, \boldsymbol{b}_i)$ found using principal component analysis (PCA).

Estimation (II)

MAP estimation

- Using vectorized versions: $\underline{Q}_v = B_v \underline{P}_v + \underline{N}_v$.
- MAP estimation for the patch, $\underline{P}_v$, given the measurement, $\underline{Q}_v$, is obtained by:

$$\hat{\underline{P}}_v(\underline{Q}_v) = \arg\max_{\underline{x}} f_{\underline{Q}_v|\underline{P}_v}(\underline{y}|\underline{x}) f_{\underline{P}_v}(\underline{x}).$$

- Estimation is given by:

$$\begin{aligned}\hat{\underline{P}}_v(\underline{Q}_v) &= (B_v^T B_v)^{-1} \cdot \left(B_v^T \underline{Q}_v - (\sigma_N^{-2} B_v^T B_v \Sigma + I)^{-1} B_v^T \underline{Q}_v \right), \\ \Sigma &= A \Sigma_H A^T + \Sigma_{\epsilon,b} + \Sigma_\epsilon.\end{aligned}$$

  - $A = (a_1, a_2, a_3) \otimes I_{N^2}$, Σ_H is the discrete 2D fBm covariance.
  - $\Sigma_{\epsilon,b}$ and Σ_ϵ are the covariance matrices of $\underline{\mathcal{E}}_b$ and $\underline{\mathcal{E}}$, respectively.

Boosting with existing algorithms

- The proposed scheme is most suitable for denoising of NST images.
- To extend the proposed algorithm for application on general images that contain structured information, we use it in tandem with existing algorithms.
- The boosted result, $\hat{X}_{boost}$, is obtained by a linear combination of our estimator, $\hat{X}_{fBm}$, and the external algorithm result, $\hat{X}_{ext}$.

$$\hat{X}_{boost} = W(\eta_1, \eta_2)\hat{X}_{ext} + (1 - W(\eta_1, \eta_2))\hat{X}_{fBm},$$

- $0 \leq W(\eta_1, \eta_2) \leq 1$ is a content-dependent weight image, set to the residual variance of the linear model of each channel.
- The linear combination allows the incorporation of any external algorithm, without adaptations of internal mechanism of the external algorithm.

Classification

Isotropic features:

- PCA-based features, used to find (a_i, b_i).
- Color model residual variances.
- H for each color channel.

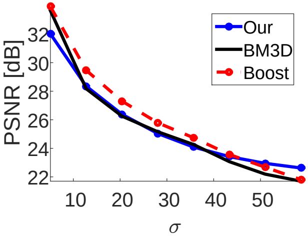
Orientation-based features:

- Gabor phase statistics: variance, kurtosis and entropy from four orientations.
- Coherence [Weickert, 1998]: $\mu = (\lambda_1 - \lambda_2)^2$, where λ_i are the eigenvalues of the smoothed structure tensor.

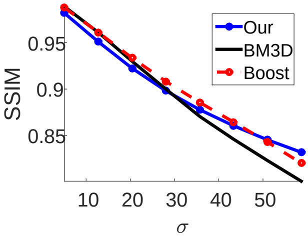
Classification method:

- Artificial neural networks (ANN)
- Training, validation and testing partition: 70%, 15% and 15%.
- Number of hidden states: average of inputs and outputs [Ripley, 1996].

Denoising results



PSNR results



SSIM results

- PSNR and SSIM results for denoising with various values of noise variances.
- Each point is the average result of all images in the dataset.
- Boosting enhances the objective performance of the MAP-based scheme.

Deblurring results

- $\sigma = 2.55$

PSNR:

	G,0.5	G,1	G,2	G,3	M,3	M,5	M,7
Noisy	32.20	24.85	21.69	20.99	26.99	24.39	22.59
Our	33.77	**27.67**	21.52	**23.07**	**29.07**	**26.84**	**25.56**
BM3D	**35.53**	26.56	**23.39**	22.24	27.11	24.45	22.67
[Fergus et al., 2006]	26.72	24.59	22.62	22.33	26.26	24.91	23.68

SSIM:

	G,0.5	G,1	G,2	G,3	M,3	M,5	M,7
Noisy	0.981	0.897	0.814	0.790	0.932	0.880	0.842
Our	0.988	**0.936**	0.789	0.842	**0.955**	**0.932**	**0.908**
BM3D	**0.989**	0.917	0.836	0.794	0.916	0.856	0.804
[Fergus et al., 2006]	0.949	0.907	**0.921**	**0.910**	0.936	0.908	0.893

- $\sigma = 7.65$

PSNR:

	G,0.5	G,1	G,2	G,3	M,3	M,5	M,7
Noisy	28.13	23.92	21.42	20.87	25.59	23.63	22.22
Our	28.66	**25.48**	22.05	**22.30**	**26.41**	**25.02**	**24.05**
BM3D	**29.21**	25.10	**22.63**	22.24	25.68	23.92	22.19
[Fergus et al., 2006]	25.98	21.16	20.88	20.45	21.16	20.78	20.59

SSIM:

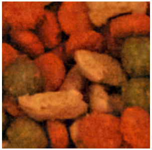
	G,0.5	G,1	G,2	G,3	M,3	M,5	M,7
Noisy	0.944	0.862	0.782	0.759	0.896	0.846	0.809
Our	**0.954**	**0.902**	0.809	**0.821**	**0.922**	**0.893**	**0.871**
BM3D	0.949	0.884	**0.820**	0.788	0.892	0.845	0.805
[Fergus et al., 2006]	0.925	0.785	0.812	0.777	0.778	0.748	0.749

Deconvolution examples

- Denoising an NST with AWGN, noise standard deviation $\sigma = 51$.
- Deblurring an NST with motion blur of size 5×5 pixels and angle of 45°, and AWGN of standard deviation $\sigma = 2.55$.

Ground truth

PSNR:15.41[dB], SSIM:0.678
Degraded image



PSNR:25.19[dB], SSIM:0.940
Our method (boosted)

PSNR:22.76[dB], SSIM:0.917
BM3D [Danielyan et al., 2012]

Ground truth

PSNR:21.57[dB], SSIM:0.870
Degraded image

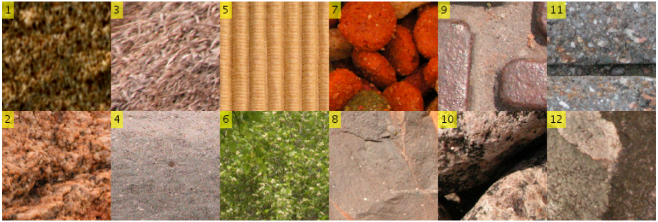
PSNR:25.39[dB], SSIM:0.940
Our method

PSNR:21.55[dB], SSIM:0.847
BM3D [Danielyan et al., 2012]

Classification results

Datasets:

- McGill [Olmos and Kingdom, 2004]: 4224 32×32-sized patches from 12, 704×512-sized images, each considered to be a different class:



- KTH-TIPS2 [Caputo et al., 2005]: 4752 64×64-sized patches (downscaled from 200×200) belonging to 11 texture classes.

⋆ High intra-class variability and complex texture structure

Testing error rates for McGill and KTH-TIPS2:

1. Using only color model features: 18.9% and 21.68%.
2. Using isotropic features (color model+residuals+fractal): 11.0% and 16.15%.
3. Using isotropic+anisotropic features: 4.1% and 7.96%.

Summary

- NST color channels are highly correlated
- An RGB color model has been derived for NST
- A linear estimation scheme shows good restoration in denoising and deblurring
- The model can be used to boost existing algorithms
- Anisotropic and isotropic features are required for NST classification

Bibliography

Benoit B Mandelbrot and John W Van Ness. Fractional Brownian motions, fractional noises and applications. *SIAM Rev.*, 10 (4):422–437, 1968.

I Zachevsky and Y Y Zeevi. Single-Image Superresolution of Natural Stochastic Textures Based on Fractional Brownian Motion. *IEEE Trans. Image Process.*, 23(5):2096–2108, 2014. ISSN 1057-7149. doi: 10.1109/TIP.2014.2312284.

Beatrice Pesquet-Popescu and Jacques L Vehel. Stochastic fractal models for image processing. *IEEE Signal Process. Mag.*, 19(5):48– 62, sep 2002. ISSN 1053-5888. doi: 10.1109/MSP.2002.1028352. URL **http://ieeexplore.ieee.org/ielx5/79/22084/01028352.pdf?tp={&}arnumber=1028352{&}isnumber=22084**.

Adriana Olmos and Frederick A A Kingdom. A biologically inspired algorithm for the recovery of shading and reflectance images. *Perception*, 33:1463–1473, 2004.

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

Ido Zachevsky and Yehoshua Y. Zeevi. Model-based Color Natural Stochastic Textures Processing and Classification. In *IEEE Glob.*, 2015.

Nir Sochen, Ron Kimmel, and Ravikanth Malladi. A general framework for low level vision. *IEEE Trans. image Process.*, 7 (3):310–8, jan 1998. ISSN 1057-7149. doi: 10.1109/83.661181. URL **http://www.ncbi.nlm.nih.gov/pubmed/18276251**.

Joachim Weickert. *Anisotropic Diffusion in Image Processing*. Teubner Stuttgart, 1998.

Brian D Ripley. *Pattern recognition and neural networks*. Cambridge university press, 1996.

Rob Fergus, Barun Singh, Aaron Hertzmann, Sam T. Roweis, and William T. Freeman. Removing camera shake from a single photograph. *ACM Trans. Graph.*, 25(3):787, 2006. ISSN 07300301. doi: 10.1145/1141911.1141956.

Aram Danielyan, Vladimir Katkovnik, and Karen Egiazarian. BM3D frames and variational image deblurring. *IEEE Trans. Image Process.*, 21(4):1715–1728, 2012.

Barbara Caputo, Eric Hayman, and P. Mallikarjuna. Class-specific material categorisation. *Proc. IEEE Int. Conf. Comput. Vis.*, II:1597–1604, 2005. ISSN 1550-5499. doi: 10.1109/ICCV.2005.54.