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]:

- 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].
The channels of color NST are correlated.

Higher correlation in small patches, better fractal behavior in larger patches.

\* \( \rho(R, G) = 0.94, \rho(R, B) = 0.84. \)
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 \( \Theta_k = (\{a_i^k, b_i^k\}_{i=1}^3, H^k) \).

Analysis performed on images from the McGill texture database [Olmos and Kingdom, 2004] and the VisTex database [Pickard et al., 1995].

- Patch size used: \( 32 \times 32 \).
- High absolute correlation is indicated between patch color channels.
Deconvolution scheme

- $Q = \{ Q_i(\eta_1, \eta_2) \}_i$: a blurred-noisy image patch.
- Degradation model:
  \[
  Q_i(\eta_1, \eta_2) = (P_i \ast B)(\eta_1, \eta_2) + N_i(\eta_1, \eta_2), \quad i \in \{1, 2, 3\},
  \]
- $B$ 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 $\sigma_{\epsilon_{b,i}}^2$, and $\{\epsilon_i(\eta_1, \eta_2)\}_i$ is an i.i.d normally distributed image with variance $\sigma_{\epsilon,i}^2$. These items allow deviations from the model by introducing model noise terms.

Two-stage deconvolution:

1. Parameter estimation, $\Theta = (\{a_i, b_i\}_i, H)$.
Parameter estimation

- The variance of the image increments for known \((\alpha, \beta)\),
  \[
  \text{var}(Q(\eta_1, \eta_2) - Q(\xi_1, \xi_2)) = \alpha \| (\eta_1, \eta_2) - (\xi_1, \xi_2) \|^2H + \beta,
  \]
  is estimated by regression of the increments' sample variance, thus obtaining \(H\) as the power exponent.

- The color model parameters, \((a_i, b_i)\), defining a patch,
  \[
  P_i(\eta_1, \eta_2) = a_i P_0(\eta_1, \eta_2) + b_i,
  \]
  are estimated using Beltrami flow [Sochen et al., 1998].

- The Beltrami flow correlates RGB image gradients by using a diffusion flow.

- Estimation of \((a_i, b_i)\) is performed on the Beltrami flow processed image.
- Optimal \((a_i, b_i)\) found using principal component analysis (PCA).
Estimation (II)

**MAP estimation**

- Using vectorized versions: $Q_v = B_v P_v + N_v$.
- MAP estimation for the patch, $P_v$, given the measurement, $Q_v$, is obtained by:

$$
\hat{P}_v(Q_v) = \arg \max_x f_{Q_v|P_v}(y|x)f_{P_v}(x).
$$

- Estimation is given by:

$$
\hat{P}_v(Q_v) = (B_v^T B_v)^{-1} \cdot (B_v^T Q_v - (\sigma_N^{-2} B_v^T B_v \Sigma + I)^{-1} B_v^T Q_v),
$$

$$
\Sigma = A\Sigma_H A^T + \Sigma_{\epsilon,b} + \Sigma_{\epsilon}.
$$

- $A = (a_1, a_2, a_3) \otimes I_{N^2}$, $\Sigma_H$ is the discrete 2D fBm covariance.
- $\Sigma_{\epsilon,b}$ and $\Sigma_{\epsilon}$ are the covariance matrices of $\mathcal{E}_b$ and $\mathcal{E}$, 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, \( \hat{X}_{\text{boost}} \), is obtained by a linear combination of our estimator, \( \hat{X}_{fBm} \), and the external algorithm result, \( \hat{X}_{\text{ext}} \).

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

- \( 0 \leq W(\eta_1, \eta_2) \leq 1 \) 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 \(\lambda_i\) 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 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:**

<table>
<thead>
<tr>
<th></th>
<th>G,0.5</th>
<th>G,1</th>
<th>G,2</th>
<th>G,3</th>
<th>M,3</th>
<th>M,5</th>
<th>M,7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Our</td>
<td>33.77</td>
<td><strong>27.67</strong></td>
<td>21.52</td>
<td><strong>23.07</strong></td>
<td><strong>29.07</strong></td>
<td><strong>26.84</strong></td>
<td><strong>25.56</strong></td>
</tr>
<tr>
<td>BM3D</td>
<td><strong>35.53</strong></td>
<td>26.56</td>
<td><strong>23.39</strong></td>
<td>22.24</td>
<td>27.11</td>
<td>24.45</td>
<td>22.67</td>
</tr>
</tbody>
</table>

**SSIM:**

<table>
<thead>
<tr>
<th></th>
<th>G,0.5</th>
<th>G,1</th>
<th>G,2</th>
<th>G,3</th>
<th>M,3</th>
<th>M,5</th>
<th>M,7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Noisy</td>
<td>0.981</td>
<td>0.897</td>
<td>0.814</td>
<td>0.790</td>
<td>0.932</td>
<td>0.880</td>
<td>0.842</td>
</tr>
<tr>
<td>Our</td>
<td>0.988</td>
<td><strong>0.936</strong></td>
<td>0.789</td>
<td>0.842</td>
<td><strong>0.955</strong></td>
<td><strong>0.932</strong></td>
<td><strong>0.908</strong></td>
</tr>
<tr>
<td>BM3D</td>
<td><strong>0.989</strong></td>
<td>0.917</td>
<td>0.836</td>
<td>0.794</td>
<td>0.916</td>
<td>0.856</td>
<td>0.804</td>
</tr>
<tr>
<td>[Fergus et al., 2006]</td>
<td>0.949</td>
<td>0.907</td>
<td><strong>0.921</strong></td>
<td><strong>0.910</strong></td>
<td>0.936</td>
<td>0.908</td>
<td>0.893</td>
</tr>
</tbody>
</table>

\[ \sigma = 7.65 \]

**PSNR:**

<table>
<thead>
<tr>
<th></th>
<th>G,0.5</th>
<th>G,1</th>
<th>G,2</th>
<th>G,3</th>
<th>M,3</th>
<th>M,5</th>
<th>M,7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Our</td>
<td>28.66</td>
<td><strong>25.48</strong></td>
<td>22.05</td>
<td><strong>22.30</strong></td>
<td><strong>26.41</strong></td>
<td><strong>25.02</strong></td>
<td><strong>24.05</strong></td>
</tr>
<tr>
<td>BM3D</td>
<td><strong>29.21</strong></td>
<td>25.10</td>
<td><strong>22.63</strong></td>
<td>22.24</td>
<td>25.68</td>
<td>23.92</td>
<td>22.19</td>
</tr>
<tr>
<td>[Fergus et al., 2006]</td>
<td>25.98</td>
<td>21.16</td>
<td>20.88</td>
<td>20.45</td>
<td>21.16</td>
<td>20.78</td>
<td>20.59</td>
</tr>
</tbody>
</table>

**SSIM:**

<table>
<thead>
<tr>
<th></th>
<th>G,0.5</th>
<th>G,1</th>
<th>G,2</th>
<th>G,3</th>
<th>M,3</th>
<th>M,5</th>
<th>M,7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Noisy</td>
<td>0.944</td>
<td>0.862</td>
<td>0.782</td>
<td>0.759</td>
<td>0.896</td>
<td>0.846</td>
<td>0.809</td>
</tr>
<tr>
<td>Our</td>
<td><strong>0.954</strong></td>
<td><strong>0.902</strong></td>
<td>0.809</td>
<td><strong>0.821</strong></td>
<td><strong>0.922</strong></td>
<td><strong>0.893</strong></td>
<td><strong>0.871</strong></td>
</tr>
<tr>
<td>BM3D</td>
<td>0.949</td>
<td><strong>0.820</strong></td>
<td>0.820</td>
<td>0.788</td>
<td>0.892</td>
<td>0.845</td>
<td>0.805</td>
</tr>
<tr>
<td>[Fergus et al., 2006]</td>
<td>0.925</td>
<td>0.785</td>
<td>0.812</td>
<td>0.777</td>
<td>0.778</td>
<td>0.748</td>
<td>0.749</td>
</tr>
</tbody>
</table>
Deconvolution examples

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

<table>
<thead>
<tr>
<th>Ground truth</th>
<th>Degraded image</th>
<th>PSNR: 15.41[dB], SSIM: 0.678</th>
<th>Our method (boosted)</th>
<th>PSNR: 25.19[dB], SSIM: 0.940</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ground truth</td>
<td>Degraded image</td>
<td>PSNR: 21.57[dB], SSIM: 0.870</td>
<td>Our method</td>
<td>PSNR: 22.76[dB], SSIM: 0.917</td>
</tr>
</tbody>
</table>

BM3D [Danielyan et al., 2012]
Classification results

Datasets:

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

- **KTH-TIPS2** [Caputo et al., 2005]: 4752 $64 \times 64$-sized patches (downscaled from 200 $\times$ 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


