Non-Local Means Denoising Using a Content-Based Search Region and Dissimilarity Kernel

Hila Berkovich, David Malah, and Meir Barzohar

ISPA 2013
8th Int'l Symposium on Image and Signal Processing and Analysis

04-Sep-13
Outline

- Introduction
  - Image Denoising
  - Standard Non-Local Means (NLM)
- Proposed NLM Modifications
- Experimental Results
- Conclusions & Future Work
Introduction: Image Denoising

- Image denoising is used to find the best estimate of the original image given its noisy version.
- Common noise model:

\[ Y = X + N, \quad N \sim \mathcal{N}(0, \sigma^2_n) \]

\( Y \) = noisy image
\( X \) = original image (unknown)
\( N \) = additive white noise

It is assumed that \( X \) and \( N \) are independent

- Patch-based denoising methods have drawn much attention.
Standard Non-Local Means (NLM)

- Introduced by Buades et. al (2005).
- Exploits image redundancy.
- **Pixel restoration**: Weighted average of all gray values within the defined search region $S_i$.

\[
\hat{X}_i = \sum_{j \in S_i} w_{i,j} Y_i
\]
Standard Non-Local Means (NLM)

Weights Definition

- The weights are based on similarity between pixel neighborhoods

\[ w_{i,k} = \frac{1}{W_i} \exp \left( -\frac{d_p(i,k)}{h^2} \right), \quad k \in S_i, \quad i \text{ is the POI} \]

\[ d_p(i,k) = \frac{1}{p^2} \left\| Y(A_i) - Y(A_k) \right\|_2^2 \]

- \( d_p(i,k) \): dissimilarity measure between neighborhoods of pixels \( i \) and \( k \)
- \( S_i \): rectangular search region of size \( M \times M \)
- \( A_i \): similarity patch of size \( p \times p \)
- \( h \): weight smoothing parameter
- \( W_i \): normalization factor \( \left( \sum_{k \in S_i} w_{i,k} \right) \)
The Parameter $h$

- The NLM algorithm is sensitive to the selection of the parameter $h$

$$w_{i,j} = \frac{1}{W_i} e^{-\frac{d_p(i,j)}{h^2}}, \ j \in S_i$$

- It is usually set to be proportional to $\sigma_n$.
- In addition, simulations suggest that $h$ should match local structure:

- There are NLM modifications that suggest to use an adaptive $h$, matched to local structure (e.g., Duval et al. 2010, Dinesh et al. 2009)

**High computational complexity**
Alternative for Using a Local $h$ – Adaptive Search Region

- **Method**: use an anisotropic **adaptive** region, which includes only pixels with similar neighborhoods to that of the POI.

- **Prior art**:
  - Gradient-based classification (Mahmoudi et al. 2005) – sensitive to noise
  - Similarity patch correlation (Dinesh et al. 2009) – a threshold is required
  - Local Polynomial Approximation combined with the Intersection of Confidence Intervals (LPA-ICI) (Sun et al. 2009) – complex and enforces contiguity of search region

creates wide edge  \(\rightarrow\) causes over-smoothing
Proposed Modification I: Adaptive Model-Based Search Region

Assumptions:

$$\forall k \in S_i^S \setminus \{i\} : X(A_i) = X(A_k) \rightarrow Y(A_i) - Y(A_k) = N(A_i) - N(A_k)$$

$$\forall j \in S_i^D : X(A_i) = C_j + X(A_j) \rightarrow Y(A_i) - Y(A_j) = C_j + N(A_i) - N(A_j)$$
Adaptive Model-Based Search Region

Distribution of Dissimilarity Measure

- A Compared patch included in $S_i^S$:

$\forall k \in S_i^S \setminus \{i\}: \frac{d_p(i, k)}{2\sigma_n^2} = \frac{1}{p^2} \left\| Y(A_i) - Y(A_k) \right\|_2^2 = \frac{1}{p^2} \sum_{m \in A_i} \sum_{l \in A_k} \left( \frac{N_m - N_l}{\sqrt{2\sigma_n}} \right)^2 \sim \chi^2_{p^2}$

$E \left[ \frac{d_p(i, k)}{2\sigma_n^2} \right] = 1, \quad Var \left[ \frac{d_p(i, k)}{2\sigma_n^2} \right] = \frac{2}{p^2}$

- A Compared patch included in $S_i^D$:

$\forall j \in S_i^D: \frac{d_p(i, j)}{2\sigma_n^2} = \frac{1}{p^2} \left\| Y(A_i) - Y(A_j) \right\|_2^2 = \frac{1}{p^2} \sum_{m \in A_i} \sum_{l \in A_k} \left( \frac{C_j + N_m - N_l}{\sqrt{2\sigma_n}} \right)^2 \sim \chi^2_{p^2} \left( \lambda_j \right)$

$\lambda_j = f(C_j)$
For $p^2 \gg 1$, the Chi-Square distribution converges to a Normal distribution.

For $p^2 = 25$
Adaptive Model-Based Search Region

Difference Between Distributions

\[ \forall k \in S_i^S \setminus \{i\}: \quad \frac{d_p(i, k)}{2\sigma_n^2} \sim \mathcal{N}\left(1, \frac{2}{p^2}\right) \]

\[ \forall j \in S_i^D: \quad \frac{d_p(i, j)}{2\sigma_n^2} \sim \mathcal{N}\left(1 + \frac{\lambda_j}{p^2}, \frac{2}{p^2} + \frac{4\lambda_j}{p^4}\right) \]

- The difference between the distributions of the two sets can serve as a classification measure.
- Since \( \lambda_j \) is unknown, we use a one-side hypothesis based on the dissimilarity variance:

**Pixels included in \( S_i^S \) are characterized by a variance \( \leq 2/p^2 \)**
Adaptive Model-Based Search Region

Classification Via Accumulated Variance

Stop accumulation once

$$Var\left\{ \frac{d_p(i,k)}{2\sigma^2_n} \right\}_{k \in S_i} > \frac{2}{p^2}$$

Compute Accumulated Variance by starting with the first 2 elements and adding one element at a time

Sort $$\left\{ \frac{d_p(i,k)}{2\sigma^2_n} \right\}_{k \in S_i}$$ in an ascending order

$$\forall k \in S_i : \frac{d_p(i,k)}{2\sigma^2_n}$$

$$S^S_i$$ $$S^P_i$$

$$d_{Th} = 6$$

Accumulated Variance

Sorted normalized $$d_p$$

$$\frac{2}{p^2}$$

$$0$$ $$2$$ $$4$$ $$6$$ $$8$$ $$10$$ $$12$$ $$14$$ $$16$$

$$(0, 2, 4, 6, 8, 10, 12, 14, 16)$$

$$(0, 5, 10, 15)$$
Examples of Adaptive Search Region of Different Local Structures
NLM with Patch–Kernel

- 2 types of patch (dissimilarity)-kernels are used frequently in NLM denoising:

\[
d_p(i,k) = \frac{1}{p^2} \| Y(A_i) - Y(A_k) \|_{2,a}^2 = \frac{1}{p^2} \sum_{m \in A_i, l \in A_k} \alpha_m (N_m - N_l)^2
\]

**Uniform patch-kernel**

**Box patch-kernel**

- Smooth regions
- Textured regions / Edges
Proposed Modification II: Patch–Kernel Type Adaptation

- The Adaptive Model-Based Search Region output provides an $S_i^S$ set per pixel.
Cluster Cardinality Map Data

- Classify the data of the normalized cardinality map using K-Means with K=2.
- The classification results in 2 centroids:
  - Large centroid value
    - Weights are computed based on Uniform patch-kernel
  - Small centroid value
    - Weights are computed based on Box patch-kernel
Patch–Kernel Type Adaptation (Cont’d)

- Cardinality map clustered data \( |S_i^S| \)
  * For \( \sigma_n = 20 \)
## Experimental Results

- **Uniform NLM vs. Adaptive NLM**

<table>
<thead>
<tr>
<th></th>
<th>Uniform</th>
<th>Adaptive</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PSNR [dB]</strong></td>
<td>24.78</td>
<td>25.62</td>
</tr>
<tr>
<td><strong>SSIM</strong></td>
<td>0.689</td>
<td>0.75</td>
</tr>
</tbody>
</table>

*For $\sigma_n = 20$

$p = 5$

$M = 11$
Experimental Results (Cont’d)

- Uniform NLM vs. Adaptive NLM

<table>
<thead>
<tr>
<th></th>
<th>Uniform</th>
<th>Adaptive</th>
</tr>
</thead>
<tbody>
<tr>
<td>PSNR [dB]</td>
<td>24.78</td>
<td>25.62</td>
</tr>
<tr>
<td>SSIM</td>
<td>0.689</td>
<td>0.75</td>
</tr>
</tbody>
</table>

* For $\sigma_n = 20$
* $p = 5$
* $M = 11$
**Experimental Results (Cont’d)**

- Box NLM vs. Adaptive NLM

<table>
<thead>
<tr>
<th></th>
<th>Box</th>
<th>Adaptive</th>
</tr>
</thead>
<tbody>
<tr>
<td>PSNR [dB]</td>
<td>25.54</td>
<td>25.62</td>
</tr>
<tr>
<td>SSIM</td>
<td>0.74</td>
<td>0.75</td>
</tr>
</tbody>
</table>

* For $\sigma_n = 20$
  
  $p = 5$
  
  $M = 11$
**Experimental Results (Cont’d)**

- **Box NLM vs. Adaptive NLM**

<table>
<thead>
<tr>
<th></th>
<th>Box</th>
<th>Adaptive</th>
</tr>
</thead>
<tbody>
<tr>
<td>PSNR [dB]</td>
<td>25.54</td>
<td>25.62</td>
</tr>
<tr>
<td>SSIM</td>
<td>0.74</td>
<td>0.75</td>
</tr>
</tbody>
</table>

* For $\sigma_n = 20$
**Experimental Results (Cont’d)**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Lena</td>
<td>20</td>
<td>30.11/0.87</td>
<td>30.27/0.86</td>
<td>30.48/0.88</td>
</tr>
<tr>
<td>Baboon</td>
<td>20</td>
<td>24.78/0.69</td>
<td>25.54/0.74</td>
<td>25.62/0.75</td>
</tr>
<tr>
<td>Barbara</td>
<td>20</td>
<td>29.11/0.87</td>
<td>29.19/0.87</td>
<td>29.33/0.88</td>
</tr>
<tr>
<td>Lena</td>
<td>30</td>
<td>28.03/0.81</td>
<td>28.03/0.78</td>
<td>28.32/0.82</td>
</tr>
<tr>
<td>Pepper</td>
<td>30</td>
<td>28.03/0.83</td>
<td>28.06/0.81</td>
<td>28.39/0.84</td>
</tr>
</tbody>
</table>
Conclusion

- Two modifications of the NLM algorithm were introduced:
  - **Model-based** adaptive search region
    - Parameter-free
    - not restricted to be contiguous
  - **Content-based** patch-kernel type
    - matched to local structure → smooth regions are less granular while texture and edges are preserved.
Future Work

• Consider correlation between dissimilarity values due to:
  - Overlap between similarity patches.
  - Patches in the same search region are compared to the same reference patch.

• Apply the suggested modifications on images characterized by Poisson noise.

**In Progress**
Non-Local Means Denoising Using a Content-Based Search Region and Dissimilarity Kernel