COMBINED NON-LOCAL AND MULTI-RESOLUTION SPARSITY PRIOR IN IMAGE RESTORATION

Jan Aelterman, Bart Goossens, Hiep Luong, Jonas De Vylder, Aleksandra Pizurica and Wilfried Philips

Ghent University - TELIN - IPI - IBBT
Sint-Pietersnieuwstraat 41, B-9000 Ghent, Belgium
jan.aelterman@telin.ugent.be

ABSTRACT

In the field of image denoising, the non-local means (NLMS) filter is a conceptually simple, yet powerful technique. This filter exploits non-local, i.e. spatially repetitive, structure in natural images to estimate noise-free structure. In contrast, a wide variety of image restoration problems have been solved exploiting local smoothness of natural images, e.g. by enforcing sparsity of images when subjected to a multi-resolution transform. In this paper we introduce the prior knowledge of non-local repetitiveness of image structures into a broad multi-resolution image restoration framework. The proposed framework allows the power of the NLMS filter, supplemented by multi-resolution sparsity, to be extended for a wide variety of image restoration problems, such as demosaicing, deconvolution, reconstruction from insufficient measurements,... in a conceptually simple way.

Index Terms— Non-local means, shearlet, sparsity, self similarity, denoising

1. INTRODUCTION

Many image restoration problems can be considered as recovery of a signal which has been corrupted by degradations that can be modeled as linear (such as blur) in combination with additive noise. Almost invariably, these inverse problems require prior knowledge, in order to regularize the problem and produce desirable results. The quality of the result is then heavily dependent on the type of prior knowledge used. Local image smoothness is a very powerful prior as natural images are typically piecewise smooth. Such regularization has many forms, such as Markov Random Field models [1], PDE-based smoothness [2] or heavy tailed multi-resolution decomposition probability densities [3]. In contrast, some algorithms exploit non-local self-similarity as an image prior, e.g. in the field of image denoising [4] and superresolution [5]. Although they are complementary, the combination of both priors has received limited attention, a notable exception is [6], where an image is considered as consisting of similar patches that can each be efficiently sparsified using a (learned) dictionary. In this paper, we present a maximum a posteriori estimator (MAP) that makes a combination of both the (non-local) self-similarity prior and the (local) multi-resolution probability density prior. The combination is valuable because of the complementary nature of the image priors. The algorithm enforces non-local self-similarity when possible, in order to avoid destroying important repetitive image details and textures. In regions where there is simply not enough self-similarity, it enforces local smoothness in a multi-resolution way.

This paper is organized as follows: In section 2, we introduce the notations and the interpretation of the proposed method as a maximum a posteriori (MAP) estimator. The non-local means (NLMS) filter is explained in terms of this MAP estimation in section 2.2. Section 3 explains the proposed method and implementation as an iterative optimization algorithm. In Section 4, the versatility of the proposed algorithm is demonstrated in some diverse image restoration scenarios. Section 5 concludes the paper.

2. BAYESIAN IMAGE RESTORATION

Consider a corruption-free image \( \bar{x} \) was sent through a linear corruption process \( \bar{H} \). The vector notation does not involve loss of generality as any N-dimensional image can be expressed as a 1D vector through column stacking and every linear corruption process can be consequently reformulated. A (white) Gaussian noise contribution \( \bar{n} \) was subsequently added. The corrupted image \( \bar{y} \) is

\[
\bar{y} = \bar{H}\bar{x} + \bar{n},
\]

where we consider the 2D images to be represented by a vector, by stacking pixels in line scanning order. A pseudo-inverse, which is the maximum likelihood estimate when it exists, leads to a highly undesirable estimate for \( \bar{x} \) since noise variance can be dramatically increased. In order to stabilize
the problem, prior knowledge is typically introduced, e.g. in the form of the classic MAP estimator:

$$\hat{x} = \arg \max_{x} \log p(\hat{x}) + \log p(\hat{y} | \hat{x})$$

(1)

with $p(\hat{x})$ the prior distribution of corruption-free images and $p(\hat{y} | \hat{x})$ the conditional distribution of corrupted images. The desirability of a certain solution is controlled through the definition of the prior knowledge term $p(\hat{x})$. The next sections will elaborate on some possible definitions.

### 2.1. Local sparsity

Natural images are highly compressible, this means that these signals can be represented accurately with relatively few coefficients (i.e. sparsely) in some well chosen basis or frame. Recently the discrete shearlet transform (DST) [7, 8] has been developed. This transform has superior approximation properties [7]. The atoms for a DST can be thought of as line structures with different orientations. A prior model based on DST sparsity will enforce a solution with as few line structures as possible; Important image aspects are retained while artifacts due to noise amplification are avoided. Such a prior model can take a simple form:

$$\log p(\hat{x}) = -\lambda_s |\tilde{S}\hat{x}|_1$$

(2)

with $|.|$ the L-1 norm and $\tilde{S}$ a sparsifying transform, such as the DST. Although this analysis based model is a non-trivial prior distribution, it is known to be a very effective model in image restoration [9] as it relates to highly kurtotic prior distributions, which are a good model for transform coefficients of natural images [10].

### 2.2. Non-local self-similarity

The non-local means denoising algorithm (NLMS) works by averaging pixels with similar neighborhoods [4]. Here, we propose to view the NLMS algorithm as the exploitation of prior information that an image (or at least some patches) is similar to translated versions of itself:

$$\log p(\hat{x}) = -\lambda_N \sum_{i,j=1}^{N} \frac{1}{\sigma_{i,j}^2} \left\| \left( \tilde{T}_i - \tilde{T}_j \right) \hat{x} \right\|_2^2$$

where $\tilde{T}_d$ selects and translates a single image pixel, where the index $d$ corresponds to a unique N-dimensional translation of image corresponding to the vector on which $\tilde{T}_d$ operates. Although again not a trivial prior distribution, it can roughly be thought of as a Gaussian model for the individual pixel differences in an image. This model becomes very powerful through control over the patch similarity parameters $\sigma_{i,j}^2$. Using this prior in the MAP estimator (1) for a denoising application (that means $\tilde{H}$ is the unit matrix), we can derive for one pixel with index $n$:

$$\hat{x}[n] = \frac{\hat{y}[n] + 2 \sum_{i}^{N} \frac{1}{\sigma_{n,i}^2} \hat{x}[i]}{1 + 2 \sum_{i}^{N} \frac{1}{\sigma_{n,i}^2}}.$$  

When the parameters $\sigma_{i,j}^2$ are chosen to correspond with the patch similarities from the NLMS denoising algorithm and when we consider preliminary estimates for the shifted pixels $\hat{x}[i] = \tilde{y}[i]$, it can be seen that this algorithm reduces to the classic NLMS denoising algorithm [4].

### 3. A COMBINED PRIOR

Both the priors discussed in sections 2.1 and 2.2 have proven their worth in natural image restoration. It should be stressed that both priors are essentially complimentary: sparsity is a local prior, as it restricts the way a signal is approximated locally (within the support of the shearlet atoms that make up a feature), while non-local self-similarity is by its very definition non-local. We propose a combined prior in the MAP image restoration problem:

$$\hat{x} = \arg \min_{x} \lambda_N \sum_{i,j=1}^{N} \frac{1}{\sigma_{i,j}^2} \left\| \left( \tilde{T}_i - \tilde{T}_j \right) \hat{x} \right\|_2^2 + \lambda_s |\tilde{S}\hat{x}|_1 + \left\| \hat{y} - \tilde{H}\hat{x} \right\|_2^2$$

(3)

We estimate the weights $\sigma_{i,j}^2$ by comparing image patches differences on a preliminary estimate of the image $\hat{x}$, using a modified bisquare function, as this increases NLMS performance [11]:

$$\hat{\sigma}_{i,j} = \begin{cases} \sqrt{2} \frac{\left\| \left( \tilde{T}_i - \tilde{T}_j \right) \hat{x} \right\|_2^2}{\left( 1 + \frac{\left\| \left( \tilde{T}_i - \tilde{T}_j \right) \hat{x} \right\|_2^2}{h^2} \right)^{\frac{3}{2}}} & \text{if } \left\| \left( \tilde{T}_i - \tilde{T}_j \right) \hat{x} \right\|_2^2 < h \\ \infty & \text{if } \left\| \left( \tilde{T}_i - \tilde{T}_j \right) \hat{x} \right\|_2^2 \geq h \end{cases}$$

(4)

where $\tilde{T}_d$ selects and translates a window of image pixels with size $w$ using an index $d$ as in the previous section. The remaining parameter $h$ controls the amount of allowed patch similarity. For pixel regions with low similarity to the rest of the image, the weights $\sigma_{i,j}^2$ are very low, hence the relative importance of the sparsity term will rise. The convex optimization problem (3) can be solved using an operator splitting approach, by introducing a new variable $d$:

$$\left\{ \hat{x}, \hat{d} \right\} = \arg \min_{x,d} \lambda_N \sum_{i,j=1}^{N} \frac{1}{\sigma_{i,j}^2} \left\| \left( \tilde{T}_i - \tilde{T}_j \right) \hat{x} \right\|_2^2 + \lambda_s |d|_1 + \left\| \hat{y} - \tilde{H}\hat{x} \right\|_2^2 \text{ s.t. } d = \tilde{S}\hat{x}$$

3050
Using an augmented Lagrangian technique [12] to solve this constrained problem, we get:

\[
\begin{align*}
\hat{x}_{l+1} &= \arg\min_{x} \lambda_N \sum_{i,j=1}^{N} \frac{1}{\sigma_i^2} \left( \left( \hat{T}_i - \hat{T}_j \right) x \right)^2 \\
&\quad + \| \hat{y} - \hat{H} x \|^2_2 + \lambda_d \| \hat{d} - \hat{S} \hat{x}_{l+1} - \hat{\mu}_k \|^2_2 \\
\hat{d}_{l+1} &= \arg\min_{d} \lambda_s |d|_1 + \lambda_d \| d - \hat{S} \hat{x}_{l+1} - \hat{\mu}_k \|^2_2 \\
\hat{\mu}_{k+1} &= \hat{\mu}_k + \left( \hat{S} \hat{x}_{l+1} - \hat{\mu}_k \right)
\end{align*}
\]

This is an algorithm with an inner and outer iteration, which needs to be iterated until both converge. In practice, it is often faster to limit the inner iteration to 1 iteration without losing significant accuracy. The outer iteration is merely an update of the augmented Lagrangian variable. The first step in the inner iteration is a linear system for the augmented Lagrangian variable. The second step in the inner iteration is a linear system for \( \hat{d} \). This can be solved very efficiently, by using a discrete Fourier transform if the matrix on the left hand side is circulant, or else by using the (bi)conjugate gradient algorithm. Memory constraints usually demand recalculation of the weights (4), this can be done within acceptable time limits using the acceleration techniques presented in [11]. The second step in the inner iteration of (5) is implemented easily, by the soft thresholding algorithm:

\[
\hat{d}_{l+1} = \text{sign} \left( \hat{S} \hat{x}_{l+1} + \hat{\mu}_k \right) \left( \hat{S} \hat{x}_{l+1} + \hat{\mu}_k - \frac{\lambda_s}{\lambda_d} \right)
\]

The algorithm usually converges within a few iterations. In order to find the optimal balance for the parameters \( \lambda_s \) and \( \lambda_N \), we use the Monte Carlo Stein unbiased risk estimator from [13]. This allows for an accurate estimate of the mean squared error without knowledge of the ground truth.

### 4. RESULTS

#### 4.1. Denoising

In a first set of experiments, we test the algorithm on a denoising experiment and compare to straightforward NLMS denoising. The proposed algorithm inherently tunes a multi-resolution local denoising in areas where there is little self-similarity. In this way, the algorithm naturally switches between the more performant non-local denoising, where possible, and the local multi-resolution denoising, where needed. We use the proposed algorithm on a version of the Lena image, where we added white Gaussian noise with standard deviation 20. We used the following parameters for the NLMS algorithm: a comparison window size \( w \) of \( 7 \times 7 \) for calculation of the weights, a limited search window of \( 15 \times 15 \) pixels and a \( h \) parameter of 38. The weight parameters were: \( \lambda_s = 15 \) and \( \lambda_N = 1 \). The same experiment was performed on the airplane image from the Kodak set, the weight parameters were: \( \lambda_s = 5 \) and \( \lambda_N = 1 \).

Some details from the resulting images for both experiments are shown in figure 1. We can see that the algorithm behaves as predicted and compares favorably with state-of-the-art sparsity enforcing methods, such as BLS-GSM [3] and straightforward NLMS. Although the proposed algorithm incorporates shearlet thresholding, it does not suffer from visually disturbing wavelet/shearlet artifacts, which manifest as spurious image structures in other multi-resolution image restoration algorithms. The largest part of the image is denoised quite well using the NLMS side of the algorithm (equivalent to \( \lambda_s = 0.0 \)), the output of which is shown in the second column. However, noise remains near important and unique details, such as the eyes of Lena. In these areas, the relative weight of the sparsity prior comes into play for the proposed algorithm and denoising performance by as much as 1dB over conventional NLMS denoising using the same parameters.

#### 4.2. Demosaicing

We can also use the proposed technique for demosaicing and denoising, which is sometimes called denoisaicing, because the (Bayer) mosaicing operation can be modeled by the linear operation \( \hat{H} \). Calculation of the weights (4) was performed based an preliminary interpolation of the mosaic image (using [14]), an idea first used for NLMS in [15]. Furthermore, as a way of sparsifying the color image space, the transition was made to the YUV color space. It is well known that the U and V components of a natural image can be approximated very accurately using a lower bandwidth, a technique also used in PAL television. For the proposed technique, the bandwidth
difference causes a significantly different variance for the shearlet coefficient distribution in the Y, U and V bands. We empirically measured this and used it to choose $\lambda_{s,Y} = 1$, $\lambda_{s,U} = 4$ and $\lambda_{s,V} = 4$ for the different Y, U and V bands so that the prior (3) becomes:

$$\log p(\vec{x}) = -\lambda_{s,Y}|S_Y \vec{x}|_1 - \lambda_{s,U}|S_U \vec{x}|_1 - \lambda_{s,V}|S_V \vec{x}|_1$$

$$-\lambda_N \sum_{i,j=1}^N \frac{1}{\sigma_{ij}} \left| \left( \vec{T}_i - \vec{T}_j \right) \vec{x} \right|_2^2.$$ 

The remaining parameters are set as in the denoising experiment. Figure 2 shows the results for an experiment with Bayer captured sensor data, corrupted by white Gaussian noise of standard deviation 20. The effect of the non-local prior in preserving repetitive features of the image can clearly be seen in the fence of the bottom row images of figure 2. It demonstrates how the non-local prior is a very valuable addition to the algorithm with just the shearlet sparsity prior enabled ($\lambda_N = 0$), shown in the third column. The increase of more than 2dB in PSNR brings it very close to state-of-the-art demosaicing+denoising algorithms.

5. CONCLUSION

The proposed combination of non-local prior with sparsity prior in an MAP image estimation framework is shown to have a value for denoising, where the current performance is very similar to state-of-the-art wavelet based denoising algorithms, such as BLS-GSM, although it suffers less from wavelet artifacts. The proposed formulation is demonstrated to naturally and smoothly switch between the non-local prior and the local prior, depending on the most applicable prior in a given spatial region. The proposed framework enables the use of non-local means in a large variety of image reconstruction problems, with demosaicing+denoising as just one example.

6. REFERENCES


