COLOR IMAGE DENOISING VIA SPARSE 3D COLLABORATIVE FILTERING WITH GROUPING CONSTRAINT IN LUMINANCE-CHROMINANCE SPACE

Kostadin Dabov, Alessandro Foi, Vladimir Katkovnik, and Karen Egiazarian

Institute of Signal Processing, Tampere University of Technology
P.O. Box 553, 33101 Tampere, Finland
firstname.lastname@tut.fi

ABSTRACT

We propose an effective color image denoising method that exploits filtering in highly sparse local 3D transform domain in each channel of a luminance-chrominance color space. For each image block in each channel, a 3D array is formed by stacking together blocks similar to it, a process that we call "grouping". The high similarity between grouped blocks in each 3D array enables a highly sparse representation of the true signal in a 3D transform domain and thus a subsequent shrinkage of the transform spectra results in effective noise attenuation. The peculiarity of the proposed method is the application of a "grouping constraint" on the chrominances by reusing exactly the same grouping as for the luminance. The results demonstrate the effectiveness of the proposed grouping constraint and show that the developed denoising algorithm achieves state-of-the-art performance in terms of both peak signal-to-noise ratio and visual quality.

Index Terms— color image denoising, adaptive grouping, block-matching, shrinkage.

1. INTRODUCTION

Recently, we proposed the block-matching and 3D filtering (BM3D) grayscale-image denoising method [1, 2]. It achieves excellent results in terms of both mean-squared-error and subjective visual quality [3]. These results, to the best of our knowledge, not only overcome those by any other prior algorithm but are a significant step beyond the current state-of-the-art in image denoising (e.g., [4]).

The BM3D is based on *grouping*, a procedure that finds mutually similar 2D image blocks and stacks them together in 3D arrays (groups). Due to the similarity between such grouped blocks, a 3D transform produces very sparse representations of the groups. It is well known that sparsity is a fundamental property allowing for effective denoising. The sparsity depends on the similarity, and thus on the grouping. Therefore, a proper grouping is crucial for the BM3D. However, measuring mutual similarity of image blocks is difficult when only noisy data are available. In this sense, grouping should be applied in a wise manner in order to achieve the desired sparsity.

In this paper we extend the grayscale BM3D method to color-image denoising. Given an *RGB* image corrupted by additive white Gaussian noise, we first transform the image to a luminance-chrominance color space (such as *YCbCr*, opponent, *YIQ*, etc.). Our generalization of the BM3D is non-trivial because we do not apply the grayscale BM3D independently on the three luminance-chrominance channels but we impose a *grouping constraint* on both chrominances. The grouping constraint means that the grouping is done only once, in the luminance (which typically has a higher SNR than the chrominances), and exactly the same grouping is reused for filtering both chrominances.

This constraint is based on the assumption that the underlying image structures (objects, edges, patterns, etc.) which determine the existence of mutually similar blocks are the same across all color channels. This assumption is typically fulfilled for natural images. Loosely speaking, we treat grouping as a color-invariant process. The constraint on the chrominances increases the stability of the grouping with respect to noise. This results in more accurate and sparsely represented groups in the chrominances and thus in more effective denoising.

2. BLOCK-MATCHING AND 3D FILTERING (BM3D)

We briefly explain the grayscale BM3D method; for a comprehensive study one can refer to [2].

Images are processed in a sliding-window manner, where for each processed window (block), used as a reference one, a 3D array is formed by stacking together blocks similar to it. We call this process "grouping" and the obtained 3D arrays we denominate "groups".

2.1. Grouping by block-matching

The grouping in BM3D is realized by block-matching. It searches in a local neighborhood for similar blocks by measuring the dissimilarity to the reference one. If the dissimilarity is smaller than a fixed threshold, the block is considered similar and is used for composing the group. The evaluation of the dissimilarity between noisy image blocks is sensitive to the noise. In particular, if the mean square difference is used as dissimilarity measure, then its variance grows quadratically with respect to the variance of the noise.

2.2. Collaborative filtering

The high similarity between grouped blocks in each 3D array enables a highly sparse representation of the true signal in a 3D transform domain. Therefore, a subsequent shrinkage of the transform spectra results in effective noise attenuation and the inverse of the 3D transform produces estimates of all blocks in the group. These estimates can be different for different blocks. Since each grouped block collaborates for the filtering of all other blocks in the group and vice versa, this filtering approach is called "collaborative filtering".
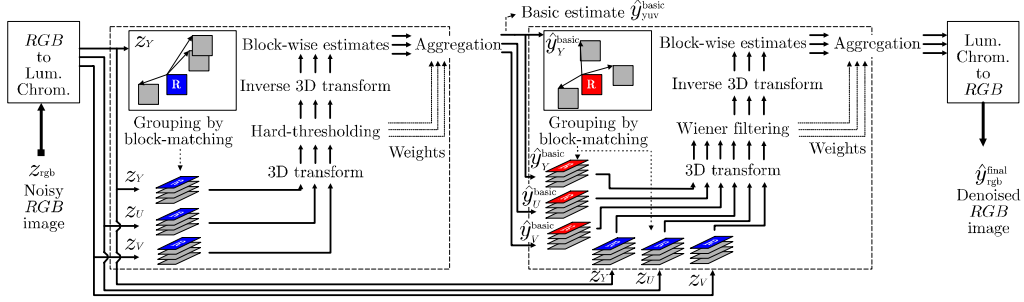
Fig. 1. Flowchart of the proposed color-image denoising method. Operations surrounded by dashed lines are repeated for each reference block.
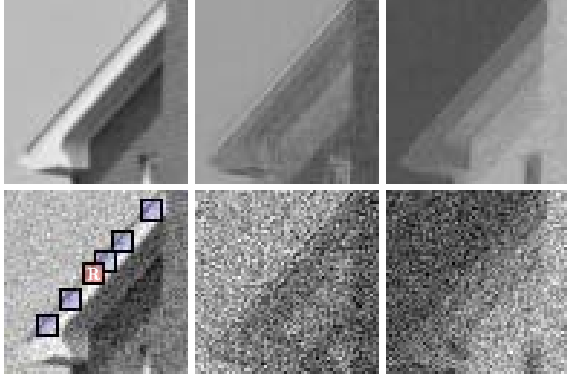

Fig. 2. The top row contains noise-free Y, Cb, and Cr channels of a fragment of *House* and the bottom row contains corresponding noisy ones (noise with standard deviation 22 is added in *RGB* space). Block-matching in the luminance Y is illustrated by showing a reference block (denoted by 'R') and the ones matched as similar to it; the same matched locations are used for both chrominances due to the grouping constraint.

3. GROUPING CONSTRAINT ON THE CHROMINANCES

Let us perform a luminance-chrominance transformation of an *RGB* natural image corrupted by independent noise, denoting the luminance with Y and the chrominances with U and V. We consider in particular the *YCbCr* and opponent color transformations, whose transform matrices are respectively

$$A_{YCbCr} = \begin{bmatrix} 0.30 & 0.59 & 0.11 \\ -0.17 & -0.33 & 0.50 \\ 0.50 & -0.42 & -0.08 \end{bmatrix}, \quad A_{opp} = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{\sqrt{6}} & 0 & \frac{-1}{\sqrt{6}} \\ \frac{1}{3\sqrt{2}} & \frac{-\sqrt{2}}{3} & \frac{1}{3\sqrt{2}} \end{bmatrix}. \quad (1)$$

Due to properties of the underlying natural color image, such as high correlation between its R, G, and B channels, the following observations can be made:

- Y has higher SNR than U and V (decorrelation of the R, G, and B channels);
- Y contains most of the valuable information (edges, shades, objects, texture patterns, etc.);
- U and V contain mostly low-frequency information (very often these channels come from undersampled data);

- iso-luminant regions with variation only in U and V are unlikely.

A naive, straightforward way to use the BM3D for color-image denoising would be to apply it separately on each of the Y, U and V channels. This approach would suffer from the lower SNR in the chrominances since the grouping is quite sensitive to the level of noise. Because a proper grouping is essential for the BM3D, we propose to perform the grouping only once for Y and reuse exactly the same grouping when applying BM3D on U and V. That is, we impose a *grouping constraint* on the chrominances based on the assumption that if the luminances of two blocks are similar, then their chrominances are also similar. An illustration of block-matching in the luminance (in *YCbCr* color space) is given in Figure 2. The observations that we made earlier about the Y, U, and V channels can be confirmed from this figure (e.g. Y has a higher SNR than both U and V, and it also contains the most valuable information about the image structures).

Furthermore, given that grouping by block-matching takes approximately half of the execution time of the BM3D, the grouping constraint enables a computational reduction of approximately one third as compared to applying the grayscale BM3D separately on Y, U and V.

4. C-BM3D ALGORITHM FOR COLOR-IMAGE DENOISING

We consider a noisy *RGB* image z_{rgb} modeled as $z_{rgb} = y_{rgb} + \eta_{rgb}$, where $y_{rgb} = [y_R, y_G, y_B]$ is the true image and $\eta_{rgb} = [\eta_R, \eta_G, \eta_B]$ is independent Gaussian noise, where $\eta_C(\cdot) \sim \mathcal{N}\left(0, \sigma_C^2\right)$ for $C = R,G,B$; the variances σ_R^2, σ_G^2, and σ_B^2 can be different. Following is the proposed algorithm, denominated Color-BM3D (C-BM3D), whose flowchart is shown in Figure 1.

1. Transform z_{rgb} to a luminance-chrominance space, the result is $z_{yuv} = [z_Y, z_U, z_V]$. The applied transform is not restricted; e.g., *YCbCr*, opponent, or others can be used.

2. *Obtain the basic estimate, denoted* $\widehat{y}_{yuv}^{basic} = \left[\widehat{y}_Y^{basic}, \widehat{y}_U^{basic}, \widehat{y}_V^{basic}\right]$.

   (a) *Block-wise estimates*. For each block in z_Y, do the following. Use block-matching to find the locations of the blocks in z_Y that are similar to the currently processed one. For each of the Y, U, and V channels, form a 3D array (group) by stacking blocks located at the obtained locations in z_Y, z_U, and z_V, respectively. Subsequently,

Method	σ	Lena (512×512 RGB)				Peppers (512×512 RGB)				Baboon (512×512 RGB)			
		10	15	20	25	10	15	20	25	10	15	20	25
C-BM3D (proposed)		**35.22**	**33.94**	**33.02**	**32.27**	**33.78**	**32.60**	**31.83**	**31.20**	30.64	**28.39**	**26.97**	**25.95**
P.wise SA-DCT [5, 6]		34.95	33.58	32.61	31.85	33.70	32.42	31.57	30.90	30.62	28.33	26.89	25.86
ProbShrink-MB [7]		34.60	33.03	31.92	31.04	33.44	32.05	31.12	30.35	30.17	27.83	26.38	25.27
VMMSE [8]		34.02	31.89	30.24	28.88	33.12	31.13	29.67	28.45	**30.68**	28.24	26.63	25.36

Table 1. Comparison of the output PSNR of a few denoising methods.

for each of these three formed groups, apply a 3D transform, attenuate the noise by hard-thresholding its transform spectrum, and invert the 3D transform to produce estimates of all grouped blocks.

(b) *Aggregation*. Compute each of $\widehat{y}_Y^{\text{basic}}$, $\widehat{y}_U^{\text{basic}}$, and $\widehat{y}_V^{\text{basic}}$ by a weighted average of the obtained block-wise estimates corresponding to the Y, U, and V channels, respectively.

3. *Obtain the final estimate* $\widehat{y}_{\text{yuv}}^{\text{final}}$, by using $\widehat{y}_{\text{yuv}}^{\text{basic}}$ to further improve the grouping and to perform collaborative Wiener filtering.

(a) *Block-wise estimates*. For each block in $\widehat{y}_Y^{\text{basic}}$, do the following. Use block-matching to find the locations of the blocks in $\widehat{y}_Y^{\text{basic}}$ that are similar to the currently processed one. For each of the Y, U, and V channels, form a pair of 3D arrays by stacking blocks located at the obtained locations in z_Y and $\widehat{y}_Y^{\text{basic}}$, z_U and $\widehat{y}_U^{\text{basic}}$, and z_V and $\widehat{y}_V^{\text{basic}}$, respectively. This results in the formation of six 3D arrays (groups). Subsequently, apply a 3D transform on each of these six 3D arrays and perform 3D Wiener filtering on the three noisy ones using the energy spectra of the ones from the basic estimate. Then, invert the 3D transform to produce estimates of all grouped blocks for each of the color channels.

(b) *Aggregation*. Compute $\widehat{y}_Y^{\text{final}}$, $\widehat{y}_U^{\text{final}}$, and $\widehat{y}_V^{\text{final}}$ by a weighted average of the obtained block-wise estimates corresponding to the Y, U, and V channels, respectively.

4. Obtain the final denoised image $\widehat{y}_{\text{rgb}}^{\text{final}}$ by transforming $\widehat{y}_{\text{yuv}}^{\text{final}}$ to RGB space.

We note that the shrinkage (hard-thresholding and Wiener filtering) uses the variance of the noise in each of the Y, U, and V channels, which are computed by $\left[\sigma_Y^2, \sigma_U^2, \sigma_V^2\right]^T = A^2 \left[\sigma_R^2, \sigma_G^2, \sigma_B^2\right]^T$, where A^2 is the used luminance-chrominance transform matrix, e.g. Eq. (1), whose elements are squared and where the superscript ^{T} denotes transposition.

5. EXPERIMENTAL RESULTS

The implementation of the C-BM3D algorithm used in the following experiments is publicly available[1] and it is based on the "Normal Profile" of the BM3D algorithm specified in [2] without any modification of its parameters. The computational complexity of the algorithm is quite reasonable; e.g., a 256×256 RGB image is denoised for about 5 seconds on a 3 GHz Pentium PC. The used test color images are *House*, *Lena*, *Peppers*, *F16*, *Baboon* and the 12th image from the Kodak PhotoCD dataset. In all experiments the

---

[1]Matlab code available at www.cs.tut.fi/~foi/GCF-BM3D.

σ	Lena	Peppers	Baboon	F16	House	Kod. 12
5	37.82	36.82	35.25	39.68	38.97	41.34
10	35.22	33.78	30.64	36.68	36.23	37.84
15	33.94	32.60	28.39	34.99	34.85	35.95
20	33.02	31.83	26.97	33.77	33.84	34.70
25	32.27	31.20	25.95	32.78	33.03	33.76
30	31.59	30.61	25.14	31.93	32.34	32.96
35	30.91	30.00	24.46	31.13	31.58	32.24
50	29.72	28.68	23.14	29.40	30.22	30.74
75	28.19	27.12	21.70	27.60	28.33	29.31

Table 2. Output PSNR of the proposed C-BM3D algorithm.

noise variance is the same for each of the R, G, and B channels, $\sigma_R^2 = \sigma_G^2 = \sigma_B^2 = \sigma^2$. We always calculate the peak signal-to-noise ratio (PSNR) in the RGB space. Unless specified otherwise, we utilize luminance-chrominance transformation to the opponent color space.

In Table 2, we present peak signal-to-noise ratio (PSNR) results of the proposed algorithm, where the PSNR is computed in the RGB space. A comparison with other methods is given in Table 1; to our knowledge, [7, 6] represent the current state-of-the-art in color-image denoising. One can see that the proposed algorithm outperforms all of the other methods for the three test images presented there, except for *Baboon* in the case of $\sigma = 10$.

Figure 3 shows a noisy ($\sigma = 35$) *House* image and the corresponding C-BM3D estimate. Most of the image details are well-preserved in the denoised image (e.g. edges, singularities, uniform areas) and in the same time there are almost no visible artifacts. The enlarged fragment in the same figure shows how effective is the proposed method for well structured details, such as the roof of the house.

6. DISCUSSION AND CONCLUSIONS

It is worth comparing the performance of the proposed C-BM3D versus the independent application of the grayscale BM3D on the individual color channels. Table 3 shows that the C-BM3D achieves $0.2 - 0.4$ dB better PSNR than the independent application of BM3D on the opponent color channels and $0.3 - 0.8$ dB better PSNR than the independent application of BM3D on the RGB channels. This improvement is not insignificant and shows the benefit of the proposed grouping constraint on the chrominances.

Although iso-luminant image regions with variation only in the chrominances are unlikely in natural images, we show such example in Figure 4 for $YCbCr$ color space. The luminance of the swimsuit is nearly constant in contrast to the Cr chrominance channel that contains a sharp transition. Note also the significantly lower SNR of the chrominance channel as compared with the luminance. Because of the imposed grouping constraint, one can expect that the sharp transition in the chrominance is not well reconstructed. However,

Fig. 3. On the left: noisy ($\sigma = 35$) *House* and a fragment of it; on the right: the corresponding denoised image (PSNR 31.58 dB) and fragment.

Approach	*Lena*	*House*	*Peppers*
C-BM3D, with grouping constraint in opponent color space	**32.27**	**33.03**	**31.20**
BM3D independently on the opponent color channels	32.01	32.64	31.01
BM3D independently on the *R*, *G*, and *B* color channels	31.44	32.18	30.93

Table 3. PSNR results of C-BM3D and of the independent application of the grayscale BM3D on the noisy color channels; the noise in *RGB* has $\sigma = 25$; all PSNR values were computed in *RGB* space.

even in such a case of grouping potentially dissimilar blocks, the proposed method can still achieve satisfactory reconstruction. The reason for this is that the BM3D uses a complete 3D transform to represent each group. The shrinkage will not be able to attenuate the noise as effectively but the chrominance details will be preserved, as shown in the denoised swimsuit in Figure 4.

We note that a similar idea of filtering the chrominances using information from the luminance was exploited already in the Pointwise SA-DCT denoising method [5, 6]. There, adaptive-shape estimation neighborhoods are determined only for *Y* and then reused for both *U* and *V*. The PSNR improvement (0.1 − 0.4 dB) of the proposed approach compared with [5, 6] is consistent with the improvement between the grayscale versions of these two methods.

The proposed C-BM3D achieves state-of-the-art performance in terms of both PSNR and subjective visual quality. This is achieved at a reasonable computational complexity. In addition, effective complexity scalability can be realized by exploiting the complexity/performance trade-off of the BM3D presented in [2].
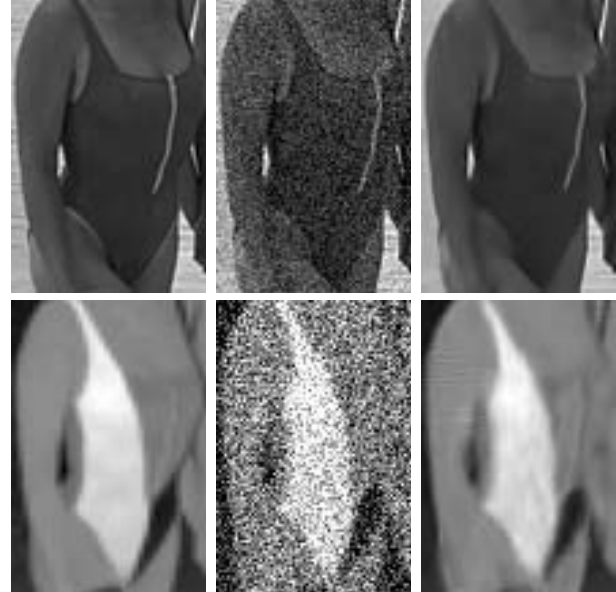

Fig. 4. Original (left), noisy (middle), and denoised (right) fragments of the luminance (top) and the chrominance *Cr* (bottom) of the 12$^{\text{th}}$ Kodak test image in *YCbCr* color space. The noise (added in *RGB* space) has $\sigma = 25$.

7. REFERENCES

[1] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising with block-matching and 3D filtering," in *Proc. SPIE Electronic Imaging: Algorithms and Systems V*, vol. 6064A-30, San Jose, CA, USA, January 2006.

[2] ——, "Image denoising by sparse 3D transform-domain collaborative filtering," *IEEE Trans. Image Process.*, December 2006, accepted, preprint at www.cs.tut.fi/~foi/GCF-BM3D.

[3] E. Vansteenkiste, D. Van der Weken, W. Philips, and E. E. Kerre, "Perceived image quality measurement of state-of-the-art noise reduction schemes," in *Lecture Notes in Computer Science ACIVS*, vol. 4179, Antwerp, Belgium, September 2006, pp. 114–124.

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

[5] A. Foi, V. Katkovnik, and K. Egiazarian, "Pointwise shape-adaptive DCT denoising with structure preservation in luminance-chrominance space," in *Proc. Video Process. Qual. Metrics Cons. Electr.*, Scottsdale, AZ, USA, January 2006.

[6] ——, "Pointwise shape-adaptive DCT for high-quality denoising and deblocking of grayscale and color images," *IEEE Trans. Image Process.*, vol. 16, no. 5, May 2007.

[7] A. Pizurica and W. Philips, "Estimating the probability of the presence of a signal of interest in multiresolution single- and multiband image denoising," *IEEE Trans. Image Process.*, vol. 15, no. 3, pp. 654–665, March 2006.

[8] P. Scheunders and J. Driesen, "Least squares inter-band denoising of color and multispectral images," in *Proc. IEEE Int. Conf. Image Process.*, Singapore, October 2004.