Structure-driven Adaptive Non-local Filter for High Efficiency Video Coding (HEVC)

Jian Zhang†, Chuanmin Jia†, Nan Zhang‡, Siwei Ma†, and Wen Gao†

†Institute of Digital Media & Cooperative Medianet Innovation Center
Peking University, Beijing, 100871, P. R. China
‡School of Biomedical Engineering, Capital Medical University, Beijing, 100069, P. R. China
{jian.zhang,cmjia,swma,wgao}@pku.edu.cn;zhangnan@ccum.edu.cn

Abstract

Deblocking filter (DF) in High Efficiency Video Coding (HEVC) is only applied to all samples adjacent to prediction units (PU), or transform units (TU), which actually exists two issues. The first one is that DF in HEVC does not fully exploit nonlocal similarity structure information in video. The second one is that DF in HEVC does not consider the inside pixels, which often suffer from quantization distortion. To alleviate these issues, in this paper, a structure-driven adaptive non-local filter (SANF) is proposed by simultaneously enforcing the intrinsic local sparsity and the non-local self-similarity of each frame. Not only SANF deals with the boundary pixels, but also the inside area, which is able to effectively reduce block artifacts while enhancing the quality of the deblocked frames. Applying SANF to luma and chroma components after DF, simulation results demonstrate that the proposed SANF can save BD-rate reduction up to 10.3% with ALF off. For luma component, SANF achieves 4.1%, 3.3%, 4.4% BD-rate saving for all intra, low delay B and random access configurations, respectively with ALF off. Furthermore, the performance with ALF on is also discussed.

1 Introduction

In April 2010, the ITU-T Video Coding Expert Group (VCEG) and ISO/IEC Moving Picture Experts Group (MPEG) formed the Joint Collaborative Team on Video Coding (JCT-VC) to develop the new coding standard, which is known as high efficiency video coding (HEVC) or H.265, formally published in 2013 [1]. Many new coding tools and coding structures are adopted in HEVC, which enable a major advance in compression relative to its predecessors.

In the HEVC video coding standard, the in-loop filters play an important role, which are applied in the encoding and decoding loops, after the inverse quantization but before saving the picture to the decoded picture buffer. HEVC standard specifies two types of in-loop filters, a deblocking filter, which is applied first, and a sample adaptive offset (SAO), which is applied to the output of the deblocking filter [2]. The purpose of the deblocking filter is to attenuate the discontinuities at prediction and transform block boundaries, while the SAO aims to further improve the quality of the decoded picture by reducing the ringing artifacts and changes in the sample intensity of areas of a reconstructed picture. Due to that the deblocking and SAO attenuate different artifacts, their benefits are usually additive when exploited together.

In HEVC, both the motion prediction and transform coding are block-based. The size of motion predicted blocks varies from 8×4 and 4×8, to 64×64 luma samples,

1068-0314/16 $31.00 © 2016 IEEE
DOI 10.1109/DCC.2016.105

while the size of block transforms and intra-predicted blocks varies from 4×4 to 32×32 samples. These blocks are coded relatively independently from the neighboring blocks and approximate the original signal with some degree of similarity. Since coded blocks only approximate the original signal, the difference between the approximations may cause discontinuities at the prediction and transform block boundaries [3] [4]. These discontinuities are attenuated by the deblocking filter.

In HEVC deblocking, the vertical boundaries in a picture are filtered first, followed by the horizontal boundaries. In a coding unit, the vertical boundaries between the coding blocks are processed starting from the left-most boundary towards the right-hand side. The horizontal boundaries are processed starting from the top-most boundary towards the bottom [4]. Since the deblocking filter (DF) is only applied to the boundaries between the coding units, prediction units, or transform units, which actually exists two issues. On one hand, the simple DF in HEVC does not fully exploit structure information in video. On the other hand, DF in HEVC does not consider the inside area, which often suffer from quantization distortion. Therefore, these two issues degrade the quality of the deblocking frames.

In the past several years, non-local self-similarity has been emerging as one of the most properties of natural image and videos, which depicts the repetitiveness of higher level patterns (e.g., textures and structures) globally positioned in images and videos, and has achieved great success in various image/video restoration applications [5–10]. In this paper, to enhance the quality of the deblocking frames, a non-local structure-based filter (SANF) is proposed by simultaneously enforcing the intrinsic local sparsity and the non-local self-similarity of each frame in video. SANF not only deals with the boundaries, but also deals with the inside areas, which is able to effectively reduce the artifacts while enhancing quality of the deblocked frames. Experimental results demonstrate that, compared with the original HEVC reference encoder implementation in all coding configuration, the proposed SANF can achieve encouraging BD-rate saving by applying SANF after DF in HEVC.

The remainder of this paper is organized as follows. Section 2 elaborates the proposed structure-driven adaptive non-local filter (SANF). Extensive experimental results are reported in Section 3. In Section 4, we concludes this paper.

2 Proposed Structure-driven Adaptive Non-local Filter

In our latest work [10], a new sparse representation model in the unit of group instead of block, named as group-based sparse representation (GSR), is proposed, which is able to exploit the local sparsity and the nonlocal self-similarity of natural images simultaneously in a unified framework. Motivated by the success of GSR in image inpainting, deblurring and compressive sensing recovery applications, in this paper, we propose to design the structure-driven adaptive non-local filter (SANF) on the basis of GSR model [10]. In this paper, different from GSR [10], the proposed SANF has the advantages of being non-iterative and being parameter-free to the QP. The details are given below.

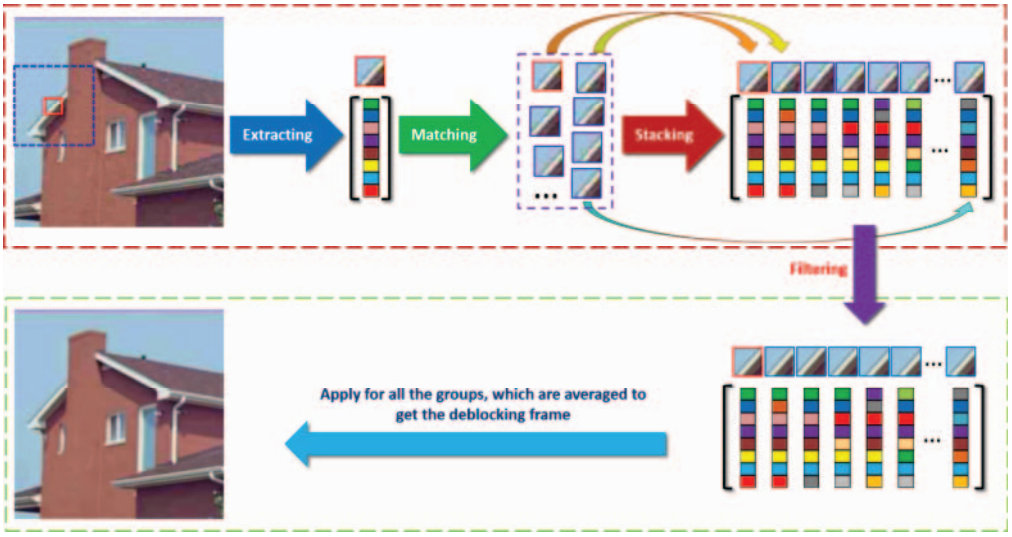

Figure 1: Illustrations for the proposed structure-driven adaptive non-local filter (SANF).

2.1 Group Construction

In this subsection, we will show how to exploit non-local structure within each frame in sequence to construct a group.

The basic idea of GSR is to adaptively sparsify the natural image in the domain of group. Thus we first show how to construct a group. In fact, each group is represented by the form of matrix, which is in fact composed of nonlocal blocks with similar structure. To be concrete, as illustrated in Fig. 1, first, divide the frame $\mathbf{x} \in \mathbb{R}^{N^2}$ with size N^2 into K overlapped blocks of size $\sqrt{B_s} \times \sqrt{B_s}$, and each block is denoted by the vector $\mathbf{x}_k \in \mathbb{R}^{B_s}$, i.e. $k = 1, 2, \ldots, K$. Then, for each block $\mathbf{x}_k$, denoted by small red square in Fig. 1, within the $W_s \times W_s$ training window (big blue square), search its c best matched blocks, which comprise the set $\mathbf{S}_{\mathbf{x}_k}$. Here, Euclidean distance is selected as the similarity criterion between different blocks. Next, all the blocks in $\mathbf{S}_{\mathbf{x}_k}$ are stacked into a matrix of size $B_s \times c$, denoted by $\mathbf{X}_{G_k}$, which includes every block in as its columns, i.e. $\mathbf{X}_{G_k}=[\mathbf{x}_{G_k \otimes 1}, \mathbf{x}_{G_k \otimes 2}, \ldots, \mathbf{x}_{G_k \otimes c}]$. The matrix $\mathbf{X}_{G_k}$ containing all the blocks with similar structure is named as a group. Note that, each block $\mathbf{x}_k$ is represented as a vector, while each group $\mathbf{X}_{G_k}$ is represented as a matrix, as shown in Fig. 1. It is obvious to observe that each block corresponds to a group.

2.2 Group-based Filtering

This subsection will give the details about how to conduct effective filter based on the group.

Given one frame $\mathbf{x} \in \mathbb{R}^{N^2}$ in the sequence, which comes from deblocking procedure, first construct K groups according to the above subsection. For each group $\mathbf{X}_{G_k}=[\mathbf{x}_{G_k \otimes 1}, \mathbf{x}_{G_k \otimes 2}, \ldots, \mathbf{x}_{G_k \otimes c}]$, $k = 1, 2, \ldots, K$, we apply singular value decom-

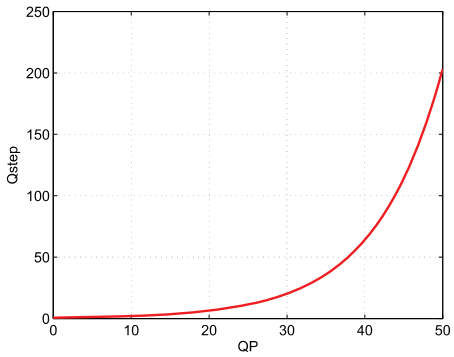

Figure 2: Relationship between quantization step (Qstep) and quantization parameter (QP).

position to it, yielding

$$\mathbf{X}_{G_k} = \mathbf{U}_{G_k}\mathbf{\Sigma}_{G_k}\mathbf{V}_{G_k}^T = \sum_{i=1}^{m} \boldsymbol{\Upsilon}_{\mathbf{x}_{G_k}\otimes i}\left(\mathbf{u}_{G_k}\otimes_i \mathbf{v}_{G_k}^T\otimes i\right), \tag{1}$$

where $\boldsymbol{\Upsilon}_{\mathbf{x}_{G_k}}$=$[\boldsymbol{\Upsilon}_{\mathbf{x}_{G_k}\otimes 1}; \boldsymbol{\Upsilon}_{\mathbf{x}_{G_k}\otimes 2}; \ldots; \boldsymbol{\Upsilon}_{\mathbf{x}_{G_k}\otimes m}]$ is a column vector, $\mathbf{\Sigma}_{G_\mathbf{k}}$=$diag(\boldsymbol{\Upsilon}_{\mathbf{x}_{G_k}})$ is a diagonal matrix with the elements of $\boldsymbol{\Upsilon}_{\mathbf{x}_{G_k}}$ on its main diagonal, and $\mathbf{u}_{G_k}\otimes i$, $\mathbf{v}_{G_k}\otimes i$ are the columns of $\mathbf{U}_{G_k}$ and $\mathbf{V}_{G_k}$, separately.

In order to reduce the artifacts and enhance the quality, the hard thresholding operation is applied to $\boldsymbol{\Upsilon}_{\mathbf{x}_{G_k}}$, i.e.,

$$\boldsymbol{\alpha}_{G_k} = hard(\boldsymbol{\Upsilon}_{\mathbf{x}_{G_k}}, \tau), \tag{2}$$

where $hard\left(\mathbf{x}, a\right) = \mathbf{x} \odot \mathbf{1}\left(abs\left(\mathbf{x}\right) - a\right)$ denotes the operator of hard thresholding and $\odot$ stands for the element-wise product of two vectors. τ denotes the threshold, whose setting is elaborated in the next subsection.

After achieving $\boldsymbol{\alpha}_{G_k}$, then we obtain the deblocking group $\hat{\mathbf{X}}_{G_k}$, expressed by

$$\hat{\mathbf{X}}_{G_k} = \sum_{i=1}^{m} \boldsymbol{\alpha}_{G_k}\otimes i\left(\mathbf{u}_{G_k}\otimes_i \mathbf{v}_{G_k}^T\otimes i\right). \tag{3}$$

This process is applied for all K groups to achieve $\hat{\mathbf{X}}_{G_k}$, $k = 1,\ 2, \ldots, K$. At last, all $\hat{\mathbf{X}}_{G_k}$ are averaged to get the filtered frame $\hat{\mathbf{x}}$.

2.3 Parameter τ Estimation

From above description, the key for SANF is the setting of parameter τ in Eq. (2). This subsection will present the details to give an adaptive and robust estimate for τ according to quantization parameter (QP) in video coding.

In this paper, we adopt Gaussian model to characterize the quantization noise between the original frame and the compressed frame due to its simplicity and effectiveness. Therefore, the estimate for τ is transformed to estimating the variance

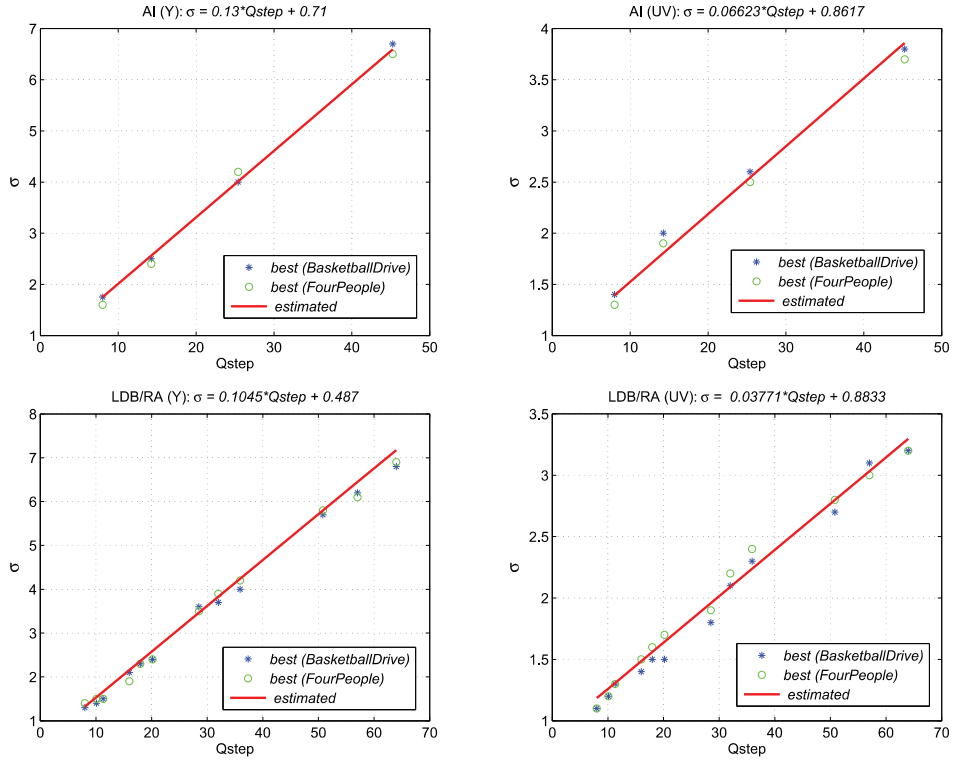

Figure 3: Agreement between the values of σ estimated by Eq. (5) and the optimal ones (found experimentally), which give the highest PSNR for the deblocked *BasketballDrive* and *FourPeople* sequences using the proposed SANF.

σ^2 of the quantization noise. Here, it is also worth emphasizing that the estimated noise variance σ^2 is not the real estimate of the variance of the difference between the original and the compressed images. Under assumption of Gaussian noise model, it is just the variance of the hypothetical Gaussian noise, which determines the level of adaptive smoothing that is able to reduce the artifacts generated by the quantization process [11].

As we know, similar to H.264/AVC, a QP is used to determine the quantization step size in HEVC. QP can take 52 values from 0 to 51 for 8-bit video sequences. An increase of 1 in QP means an increase of the quantization step size (Qstep) by approximately 12%. The resulting relationship between QP and Qstep for an orthonormal transform is formuated by [1]:

$$Qstep = 2^{\frac{(QP-4)}{6}}. \tag{4}$$

Fig. 2 also shows how Qstep increases non-linearly with QP. Since quantization consists of division by the Qstep and subsequent rounding while inverse quantization consists of multiplication by Qstep, we directly study the relationship between σ and Qstep.

First, we observe the optimal values of σ found experimentally for the sequences *BasketballDrive* and *FourPeople* compressed with different QPs (QP = 22, 27, 32,

Table 1: Coefficients in Eq. (5) for estimate σ in all configurations

Component	AI		LDB		RA	
Type	*alpha*	*beta*	*alpha*	*beta*	*alpha*	*beta*
Y	0.13	0.71	0.1045	0.487	0.1045	0.487
U	0.06623	0.8617	0.03771	0.8833	0.03771	0.8833
V	0.06623	0.8617	0.03771	0.8833	0.03771	0.8833

37) corresponding to different Qsteps calculated by Eq. (4), as illustrated in Fig. 3. It can be inferred that different sequences with the same Qp or Qstep have similar optimal values of σ, which means that σ is only related with QP or Qstep. Hence, in this paper, we propose to estimate the optimal value of σ directly from Qstep by curve fitting using the following empirical formulation:

$$\sigma = \alpha * Qstep + \beta. \tag{5}$$

The values of the two coefficients in the linear model (5), i.e. α and β, are shown in Table 1. Note that LDB configuration and RA configuration share the same coefficients. What's more, in each configuration, U component and V component utilize the same coefficients.

The red curve in Fig. 3 expresses the relationship between the estimated σ and Qstep obviously. The value τ is also has a linear relationship with σ, expressed as

$$\tau = \sigma * (B_s + \sqrt{c}). \tag{6}$$

Finally, combining Eqs. (4) (5) (6), we further obtain the relationship between the estimated τ and QP below

$$\tau = (\alpha * 2^{\frac{(QP-4)}{6}} + \beta) * (B_s + \sqrt{c}). \tag{7}$$

All samples processed by SANF would be further filtered by SAO and ALF (optional) [12]. We compare SANF result with the original DF in terms of PSNR, choosing the better one at encoder side and send flag to decoder. Extensive experiments in Section 3 will verify the robustness and effectiveness of the proposed SANF.

3 Experimental Results

This section provides simulation test data to illustrate the benefits of structure-driven adaptive non-local filter (SANF) with an analysis of the algorithm's complexity in both encoder and decoder side.

3.1 Test Conditions

Following are the experiment results conducted under the common test conditions (CTC) [13] of HEVC, which consists of six different classes (from class A to class F)

of resolutions. The resolutions of each classes are: class A is 2560×1600, class B is 1920×1080 class C is 832×480, class D is 416×240, class E is 1280×720. From class A to class E, all sequences in these classes are natural videos. The sequences in class F are not natural videos but for screen content coding (SCC) containing three resolutions: 1280×720, 1024×768, 832×480. We run our experiments from class B to class F.

- Four utilized quantization parameter (QP) values are 22, 27, 32 and 37.
- Result using Bjontegaard's method [14] in terms of BD-rate.
- Three different types of coding configurations are: all intra (AI), low delay B (LDB), random access (RA).
- All three reference configurations are tested under main conditions (Main), NOT under high efficient 10-bit conditions (Main10).
- The version of HEVC reference software [15] is HM12.0 (Anchor).
- For more persuasive analysis, ALF [16] has been added to the reference software. BD-rate performance comparison was made when ALF turns on and off respectively.

3.2 Objective Quality

Table 2 provides the PSNR results sequence-by-sequence. It is clear to observe that SANF performs well on all test sequences under all configurations. Table 3 shows us the performance of the proposed SANF under AI, LDB and RA configurations with ALF off. Obviously, SANF achieves remarkable coding gains in all circumstances. In particular, under AI configuration, SANF obtains 4.1%, 4.9%, 6,1% BD-rate reduction in YUV components, respectively. In LDB configuration, there are 3.3% BD-rate saving in Y component, and 4.4% and 5.1% BD-rate reduction in UV components. In RA configuration, SANF obtains 4.4% 5.6% and 6.1% BD-rate reduction in YUV components, respectively.

Table 4 further presents the corresponding simulation test performance with ALF on. Although ALF has been removed from TMuC in the final draft of HEVC standard (HM), its power can't be ignored. From Table 4, it is obvious to see that the proposed SANF still acquires convincing result in all configurations. The coding gains in AI are 3.1%, 2.8%, and 3.9% for YUV components, respectively. In LDB configuration, SANF obtains 2.8%, 3.7% and 4.1% for YUV components. As for RA, the coding gains are 3.4%, 4.6%, and 5.1% for each component, separately.

Although the improvement of the proposed SANF with ALF on is not as much as the improvement of the proposed SANF with ALF off, the proposed SANF with ALF on still outperforms HEVC with ALF, which fully verifies that the nonlocal structure property can obviously reduce the block artifact and the quantization noise. That also indicates, there exists extensive prospect for post filter in the next generation coding standard to explore.

Table 2: Experimental results of SANF sequence by sequence, Anchor: HM12.0 with ALF off

Sequences		AI			LDP			RA		
		Y	U	V	Y	U	V	Y	U	V
Class B	Kimono	-3.8%	-0.2%	-0.2%	-4.5%	-7.0%	-3.7%	-3.6%	-4.4%	-1.3%
	ParkScene	-2.2%	-0.1%	-0.0%	-0.4%	-0.0%	-0.2%	-1.8%	-0.8%	-0.3%
	Cactus	-3.6%	-1.4%	-7.0%	-3.5%	-3.3%	-4.5%	-5.0%	-5.8%	-5.9%
	BasketballDrive	-3.2%	-7.3%	-9.3%	-2.7%	-6.2%	-9.0%	-2.9%	-5.6%	-8.4%
	BQTerrace	-3.3%	-2.2%	-2.8%	-4.9%	-1.3%	-1.2%	-8.2%	-3.1%	-3.0%
Class C	BasketballDrill	-5.0%	-8.7%	-11.0%	-1.2%	-5.1%	-6.0%	-2.1%	-7.2%	-9.3%
	BQMall	-4.9%	-4.8%	-6.3%	-3.3%	-4.1%	-5.1%	-3.9%	-4.4%	-4.8%
	PartyScene	-2.0%	-2.2%	-3.0%	-0.2%	-0.3%	-0.4%	-0.7%	-1.0%	-0.9%
	RaceHorsesC	-2.6%	-4.0%	-7.2%	-3.0%	-4.0%	-7.6%	-3.2%	-4.2%	-8.7%
Class D	BasketballPass	-4.1%	-5.6%	-6.1%	-1.7%	-3.1%	-2.6%	-2.4%	-4.4%	-3.7%
	BQSquare	-2.0%	-0.7%	-3.8%	-0.7%	-0.6%	-1.8%	-1.3%	-0.9%	-1.4%
	BlowingBubbles	-2.5%	-5.1%	-6.2%	-0.9%	-2.9%	-2.2%	-1.5%	-3.5%	-2.4%
	RaceHorses	-3.7%	-6.2%	-7.4%	-2.9%	-3.3%	-4.8%	-3.1%	-5.0%	-6.3%
Class E	FourPeople	-6.8%	-6.8%	-7.4%	-7.7%	-8.2%	-9.1%	-9.6%	-10.0%	-10.5%
	Johnny	-7.7%	-10.2%	-9.9%	-7.2%	-12.1%	-11.4%	-10.3%	-13.6%	-13.4%
	KristenAndSara	-6.6%	-8.6%	-9.3%	-7.4%	-8.2%	-10.9%	10.0%	-10.8%	-12.3%
Class F	BasketballDrillText	-4.8%	-7.9%	-9,4%	-1.3%	-4.8%	-4.8%	-2.3%	-7.0%	-8.2%
	ChinaSpeed	-1.8%	-3.1%	-3.1%	-1.6%	-0.5%	-2.3%	-2.2%	-4.0%	-4.1%
	SlideEditing	-2.2%	-0.0%	-0.7%	-2.0%	-0.2%	-0.3%	-2.6%	-0.2%	-0.4%
	SlideShow	-5.0%	-6.4%	-6.7%	-4.7%	-5.8%	-6.9%	-4.9%	-8.7%	-8.8%
Overall		-4.1%	-4.9%	-6.1%	-3.3%	-4.4%	-5.1%	-4.4%	-5.6%	-6.1%

Table 3: Experimental results of SANF, Anchor: HM12.0 with ALF off

Sequences	AI			LDB			RA		
	Y	U	V	Y	U	V	Y	U	V
Class B	-3.2%	-2.2%	-3.9%	-3.2%	-3.5%	-3.7%	-4.3%	-3.9%	-3.8%
Class C	-3.6%	-4.9%	-6.9%	-1.9%	-3.4%	-4.8%	-2.5%	-4.2%	-5.9%
Class D	-3.1%	-4.4%	-5.9%	-1.5%	-2.5%	-2.8%	-2.1%	-3.4%	-3.4%
Class E	-7.1%	-8.5%	-8.9%	-7.4%	-9.5%	-10.5%	-10.0%	-11.4%	-12.1%
Class F	-3.5%	-4.4%	-5.0%	-2.4%	-2.8%	-3.6%	-3.0%	-5.0%	-5.4%
Overall	-4.1%	-4.9%	-6.1%	-3.3%	-4.4%	-5.1%	-4.4%	-5.6%	-6.1%

Table 4: Experimental results of SANF, Anchor: HM12.0 with ALF on

Sequences	AI			LDB			RA		
	Y	U	V	Y	U	V	Y	U	V
Class B	-1.9%	-1.0%	-2.5%	-2.1%	-2.9%	-3.3%	-2.6%	-3.0%	-3.8%
Class C	-3.1%	-2.6%	-5.0%	-2.0%	-4.0%	-5.1%	-2.2%	-4.3%	-5.9%
Class D	-2.6%	-1.6%	-3.1%	-1.6%	-2.5%	-3.0%	-1.9%	-3.6%	-3.8%
Class E	-4.9%	-4.5%	-3.9%	-5.5%	-5.5%	-5.6%	-7.5%	-7.5%	-6.8%
Class F	-3.1%	-4.3%	-5.0%	-2.8%	-3.5%	-3.6%	-2.9%	-4.7%	-5.3%
Overall	-3.1%	-2.8%	-3.9%	-2.8%	-3.7%	-4.1%	-3.4%	-4.6%	-5.1%

3.3 Subjective Quality

We further present the subjective quality in Fig. 4, which shows the reconstructed frames from sequence *Johnny* and sequence *KristenAndSara*. Obviously, under the cases of with ALF off or with ALF on, the proposed SANF not only processes the pixels near the boundary of transform units and prediction units, but also filters the

(a) (b) (c) (d)

Figure 4: Visual comparisons between anchor and proposed SANF of frames from sequences *Johnny* and *KristenAndSara*. (a) HM12.0 (LDB), ALF-OFF, QP = 37, PSNR = 36.42 dB (b) SANF (LDB), ALF-OFF, QP = 37, PSNR = 36.64 dB (c) HM12.0 (AI), ALF-ON, QP = 32, PSNR = 39.65 dB (d) SANF (AI), ALF-ON, QP = 32, PSNR = 40.07 dB. Obviously, SANF has much better quality than anchor in PSNR and visual quality.

pixels inside these blocks, providing better results on both edge and texture areas.

Finally, as for the computation complexity, nonlocal similarity property requires overlapped block-matching search in search window for each block of each frame. Therefore, the proposed algorithm can be regarded as a picture-level in-loop filter in video coding process. The encoding computation complexities of the proposed SANF in three configurations are 289% (AI), 126% (LDB) and 148% (RA) against the anchor encoder respectively with ALF off. When ALF turns on, the encoding computation complexities becomes 219% (AI), 131% (LDB) and 136% (RA) against the anchor encoder in HM12.0 with ALF on. As for the complexity of decoder side, the decoding computation complexities in all configurations take more than ten-times those of the HM12.0 decoder with ALF on/off.

4 Conclusion

In this paper, to improve the in-loop filter performance, a structure-driven adaptive non-local filter (SANF) is proposed by simultaneously enforcing the intrinsic local sparsity and the non-local self-similarity of each frame in video. SANF not only deals with the boundaries, but also inside area, which is able to effectively reduce the artifacts while enhance quality of the deblocked frames. Experimental results demonstrate that, the proposed SANF surpasses HEVC reference software convincingly. Although with high complexity, the proposed SANF shed light on the important role of post processing filter especially for future video coding standard. On going work is to address the algorithm parallelization and optimization for accelerating SANF.

5 Acknowledgement

This work was supported in part by the National Basic Research Program of China (973 Program, 2015CB351800), National Natural Science Foundation of China (No. 61572047, No. 61272255 and No. 61571017), the Postdoctoral Science Foundation of China (2015M580018),

the Scientific Research Common Program of Beijing Municipal Commission of Education (KM201310025009) and Shenzhen Peacock Plan, which are gratefully acknowledged.

6 References

[1] G. J. Sullivan, J.-R. Ohm, W.-J. Han, and T. Wiegand, "Overview of the high efficiency video coding (HEVC) standard," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 22, no. 12, pp. 1649–1668, 2012.

[2] C.-M. Fu, E. Alshina, A. Alshin, Y.-W. Huang, C.-Y. Chen, C.-Y. Tsai, C.-W. Hsu, S.-M. Lei, J.-H. Park, and W.-J. Han, "Sample adaptive offset in the hevc standard," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 22, no. 12, pp. 1755–1764, 2012.

[3] P. List, A. Joch, J. Lainema, G. Bjontegaard, and M. Karczewicz, "Adaptive deblocking filter," *IEEE transactions on circuits and systems for video technology*, vol. 13, no. 7, pp. 614–619, 2003.

[4] A. Norkin, G. Bjontegaard, A. Fuldseth, M. Narroschke, M. Ikeda, K. Andersson, M. Zhou, and G. Van der Auwera, "Hevc deblocking filter," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 22, no. 12, pp. 1746–1754, 2012.

[5] A. Buades, B. Coll, and J.-M. Morel, "A non-local algorithm for image denoising," in *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, vol. 2. IEEE, 2005, pp. 60–65.

[6] J. Zhang, R. Xiong, S. Ma, and D. Zhao, "High-quality image restoration from partial random samples in spatial domain," in *Visual Communications and Image Processing (VCIP), 2011 IEEE*. IEEE, 2011, pp. 1–4.

[7] J. Zhang, R. Xiong, C. Zhao, S. Ma, and D. Zhao, "Exploiting image local and non-local consistency for mixed gaussian-impulse noise removal," in *Multimedia and Expo (ICME), 2012 IEEE International Conference on*. IEEE, 2012, pp. 592–597.

[8] J. Zhang, D. Zhao, C. Zhao, R. Xiong, S. Ma, and W. Gao, "Image compressive sensing recovery via collaborative sparsity," *Emerging and Selected Topics in Circuits and Systems, IEEE Journal on*, vol. 2, no. 3, pp. 380–391, 2012.

[9] J. Zhang, D. Zhao, R. Xiong, S. Ma, and W. Gao, "Image restoration using joint statistical modeling in a space-transform domain," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 24, no. 6, pp. 915–928, 2014.

[10] J. Zhang, D. Zhao, and W. Gao, "Group-based sparse representation for image restoration," *Image Processing, IEEE Transactions on*, vol. 23, no. 8, pp. 3336–3351, 2014.

[11] A. Foi, V. Katkovnik, and K. Egiazarian, "Pointwise shape-adaptive dct for high-quality denoising and deblocking of grayscale and color images," *Image Processing, IEEE Transactions on*, vol. 16, no. 5, pp. 1395–1411, 2007.

[12] X. Zhang, R. Xiong, S. Ma, and W. Gao, "Adaptive loop filter with temporal prediction," in *Picture Coding Symposium (PCS), 2012*. IEEE, 2012, pp. 437–440.

[13] F. Bossen and H. Common, "Test conditions and software reference configurations, JCT-VC Doc," *L1100, Jan*, 2013.

[14] G. Bjontegaard, "Calcuation of average psnr differences between rd-curves," *Doc. VCEG-M33 ITU-T Q6/16, Austin, TX, USA, 2-4 April 2001*, 2001.

[15] F. Bossen, D. Flynn, and K. Sühring, "HEVC reference software manual," *JCTVC-D404, Daegu, Korea*, 2011.

[16] C.-Y. Tsai, C.-Y. Chen, T. Yamakage, I. S. Chong, Y.-W. Huang, C.-M. Fu, T. Itoh, T. Watanabe, T. Chujoh, M. Karczewicz *et al.*, "Adaptive loop filtering for video coding," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 7, no. 6, pp. 934–945, 2013.