Largest Coding Unit Based Framework for Non-local Means Filter

Masaaki Matsumura, Seishi Takamura and Atsushi Shimizu
NTT Media Intelligence Laboratories, NTT Corporation, Japan
E-mail: { matsumura.masaaki, takamura.seishi, shimizu.atsushi }@lab.ntt.co.jp  Tel: +81-46-859-3521

Abstract— One of the important factors for in-loop filter of video codec is low-delay capability for encoding and decoding. In this paper, we employ non-local means filter between sample adaptive offset and adaptive loop filter to the reference software of High Efficiency Video Coding HM7.0, and propose largest coding unit (LCU) based framework for non-local means filter that can reconstruct a decoded picture in LCU order at encoder and decoder. As the result, compared to HM7.0 anchor, in the case of picture-based RD-optimization, the average improvements of BD-rate for luma and chroma are 0.36 to 1.52% and 0.04 to 1.37%, respectively. Similarly, LCU-based one improves 0.20 to 1.27% and 0.67 to 1.91%, respectively. We confirm the maximum gain in the sequence of "Kimono" on low-delay P; the gains are 3.50% (Y), 2.89% (U) and 1.84% (V), respectively. Subjective quality improvements are also observed.

I. INTRODUCTION

In-loop filter that provides the reduction of visible distortions is one of important coding tools for improving coding efficiency and visual quality. Reference software of High Efficiency Video Coding (HEVC) HM7.0 [1] employs two kinds of in-loop filters, deblocking filter and sample adaptive offset (SAO).

Deblocking filter smoothes the boundaries of each coding unit, and contributes the reduction of visible blocky noises. SAO modifies each signal of a decoded picture by additional offsets according to the feature of the textural edges or the cluster of signal of the pixel. However, HEVC does not support reduction of mosquito noises that appears around high frequency domains.

We utilize non-local means (NLM) filter [2] as one of the in-loop filters of HEVC to reduce the mosquito noises, and propose LCU-based framework for NLM filter that can reconstruct decoded picture in LCU order at encoder/decoder to maintain zero-delay capability.

II. PREVIOUS WORKS

In the research for image processing, many filters such as Gaussian filter [3], bilateral filter [4] and anisotropic filter [5] are proposed to reduce the random noises for digital images. In such filters, NLM filter proposed by Buades et al. provides favorable results maintaining textural edges. NLM filter derives the filter coefficients of each element of window on-the-fly by exponential function according to filtering degree $\sigma$ and sum of squared difference between templates illustrated in Figure 1. Let $v(x, y)$ denote the pixel value at a coordinate $(x, y)$ of the decoded picture, and let variables $k$ and $l$ denote the relative coordinate corresponding to template $T$. The filter coefficient $w_{ij}(x, y)$ with the relative coordinate $(i, j)$ corresponding to window $\Omega$ is derived by

$$w_{ij}(x, y) = e^{-\frac{\sum_{l \in \Omega} [v(x+i,k,y+l)+v(x+k,y+l)]^2}{\sigma^2}} \quad (1).$$

Then filtered output $\text{PicFlt}(x, y)$ is derived by

$$\text{PicFlt}(x, y) = \frac{\sum_{(i,j) \in \Omega} w_{ij}(x, y)v(x+i, y+j)}{\sum_{(i,j) \in \Omega} w_{ij}(x, y)} \quad (2).$$

Since NLM filter smoothes the areas with similar geometric structures, low-to-middle filtering degree will reduce mosquito noises by smoothing of maintaining the textural edges, and middle-to-high filtering degree will reduce also blocky noises. Therefore, visible improvements of subjective picture quality are reported in [6-9].

However, these works permit the pixel reference between the LCU belonging to the target pixel to be filtered and other LCUs. Since HEVC encodes/decodes each LCU in raster scan order, for the decoder, the pixel reference to un-reconstructed LCUs such as lower LCU and right neighboring LCU causes the delay of reconstruction for decoded picture. Furthermore, since these works generate filtering degree picture by picture, for the decision of the filtering degree, these encoders have to wait for the reconstruction of decoded picture.

In this paper, we utilize NLM filter after SAO illustrated in Figure 2, and propose LCU-based framework for NLM filter that can reconstruct decoded picture in LCU order at encoder/decoder to maintain zero-delay capability.
In this section, proposed NLM filter is described. At first, the generation of a half-shrunk decoded picture is described in sub-clause III-A; after that, the setting for template and window are described in sub-clause III-B and III-C. Finally, the structure for LCU-based encoding and decoding containing boundary settings and syntax definition are described in sub-clause III-D and III-E.

A. Generation of a half-shrunk decoded picture

In our approach, both template shape and window shape are modified at every 2x2 sub-block according to textural feature of a half-shrunk decoded picture $\text{PicDecHalf}$ with 2-D coordinate defined by

$$
\text{PicDecHalf}_{x, y} = \text{PicDec}_{x, y} + \text{PicDec}_{x+1, y} + \text{PicDec}_{x, y+1} + \text{PicDec}_{x+1, y+1}
$$

(3),

where $x_2$ and $y_2$ are the coordinate in $\text{PicDecHalf}$ defined as $x_2 = x / 2$ and $y_2 = y / 2$.

B. Template settings

Single pixel template and Cross-shaped template (luma component only) illustrated in Figure 3 are utilized as the template of proposed NLM filter. We decide a template from two templates according to the activity in $\text{PicDecHalf}$. Gradients $dx$, $dy$ and Activity around $\text{PicDecHalf}_{x_2, y_2}$ are derived by

$$
dx = \text{PicDecHalf}_{x_2+1, y_2} - \text{PicDecHalf}_{x_2, y_2}
$$

(4),

$$
dy = \text{PicDecHalf}_{x_2, y_2+1} - \text{PicDecHalf}_{x_2, y_2}
$$

(5),

$$
\text{Activity} = (|dx| + |dy|) / \text{BitIncrement}
$$

(6),

where $\text{BitIncrement}$ is the number of additional bit for internal bit depth increase (in main profile of HEVC, $\text{BitIncrement}$ is set to 0).

C. Window settings

In the research for image processing, square-shaped window such as 7x7, 9x9 or more is generally employed. However, excessive allocation of window size causes the increase of computational complexity with tiny improvements of coding gains. Therefore, in the case of chroma component, window shape is set in constant checker-board shape illustrated in Figure 5. In the case of luma component, the window shape is allocated from 10 patterns (illustrated in Figure 6) corresponding to the index of $\text{DirectionIdx}$ adaptively, according to the direction of the textural edge $\text{Dir}$. The textural edge $\text{Dir}$ and the index of $\text{DirectionIdx}$ are derived by

$$
\text{Dir} = 2^{10} \frac{dx}{dy}
$$

(7),

$$
\text{DirectionIdx} =
\begin{cases}
5 & \text{if } dy \text{ is equal to } 0 \\
6 & \text{otherwise, if } \text{Dir} < -5120 \\
7 & \text{otherwise, if } \text{Dir} < -1707 \\
8 & \text{otherwise, if } \text{Dir} < -1024 \\
9 & \text{otherwise, if } \text{Dir} < -614 \\
0 & \text{otherwise, if } \text{Dir} < 205 \\
1 & \text{otherwise, if } \text{Dir} < 614 \\
2 & \text{otherwise, if } \text{Dir} < 1024 \\
3 & \text{otherwise, if } \text{Dir} < 1707 \\
4 & \text{otherwise, if } \text{Dir} < 5120 \\
5 & \text{otherwise }
\end{cases}
$$

(8).
D. Boundary settings

To avoid the pixel reference to un-reconstructed LCUs such as lower LCUs and right neighboring LCU, the pixels allocated at the bottom 2-pixel lines and the right 2-pixel lines of each LCU are not filtered. In the case of luma component, to 2×2 sub-blocks with cross-shaped template allocated at 3 to 4-pixel lines from the LCU boundary, the template is swapped to single pixel template.

To LCUs with the slice boundary or tile boundary, the pixels allocated at the 2-pixel lines from these boundaries are not filtered. In the case of luma component, to 2×2 sub-blocks with cross-shaped template allocated at 3 to 4-pixel lines from these boundaries, the template is swapped to single pixel template. Figure 7 illustrates the boundary settings.

E. Syntax definition

Syntax for proposed NLM filter is added to slice header and each LCU. Additional syntax element signaled in slice header is NLM filter control flags for each component. When the flag is set to 0, the component is not filtered. If picture-based rate-distortion (RD) optimization controlling NLM filter on or off is required, the switching of this flag cannot apply in low-delay encoding. For LCU-based RD-optimization, these flags are set always to 1.

When the flag is set to 1, additional syntax elements signaled in each LCU are defined as:

- Filter control flag for LCU
- Filter type
- Filtering degree
- Quad-tree control flag (luma component only)

When the filter control flag for LCU is set to 0, the LCU is not filtered. Filter type declares the reference type that decides the filtering degree of the LCU. We prepare 4 types MERGE_LEFT, MERGE_UP, SET_AVG and GEN PARAM. If the type is set to MERGE_LEFT or MERGE_UP, the filtering degree is quoted from left or upper LCU. Otherwise, if the type is set to SET_AVG, the filtering degree is set to the mean value of the filtering degrees of all filtered LCU. Otherwise, the filtering degree is signaled. For luma component, the LCU is split into quad-tree based sub-blocks; the split flags and filter enable/disable control flags of sub-block are signaled.

IV. EXPERIMENTAL RESULT

We used 23 test sequences described in Table 1 and compared our proposal with HM7.0 anchors. As the configurations for NLM filter, thresholds $Th_{LP}$, $Th_{LX}$ and $Th_C$ are set to 4, 50 and 4, respectively. The configurations for HM7.0 were set to:

- GOP: All Intra, Random Access, Low-delay B, Low-delay P
- Profile: Main profile
- Number of frames to be encoded: 60
- QP: 22, 27, 32, 37
- Number of slice: 1
- Multiple tiles/Wavefront parallel processing: Disabled

A. Objective evaluation

Coding efficiencies calculated by BD-rate [10] and computational complexities compared to the anchor are presented in Table 2 (picture-based RD-optimization) and Table 3 (LCU-based RD-optimization).

In the case of picture-based RD-optimization, the average BD-rate for luma and chroma improved 0.36 to 1.52% and 0.04 to 1.37% respectively. In the case of LCU-based one, the average BD-rate for luma and chroma improved 0.20 to 1.27% and 0.67 to 1.91% respectively. In the case of LCU-based RD-optimization, since NLM filter cannot switch NLM filter control flag in the slice header, the average BD-rate of luma lose in all GOP compared to the result of picture-based one. Especially, the differences of inter slice are greater than intra slice. However, the average BD-rate of chroma is improved.
B. Subjective evaluation

Figure 8 and 9 demonstrate the frames where subjective picture quality differences are distinguished. In both cases, mosquito noises around marked edges are more visible in the anchor than proposal.

V. CONCLUSIONS

In this paper, we utilized NLM filter after SAO to HM7.0, and proposed LCU-based framework for NLM filter that can reconstruct decoded picture in LCU order at encoder/decoder. As the result, compared to HM7.0 anchor, in the case of picture-based RD-optimization, the average improvements of BD-rate for luma and chroma were 0.36 to 1.52% and 0.04 to 1.25%, respectively. We confirmed the maximum gain in the sequence of “Kimono” on low-delay P; the gains were 3.50% (Y), 2.89% (U) and 1.84% (V), respectively. Subjective quality improvements were also observed.

REFERENCES


Table 2: Experimental results for picture-based RD-optimization.

<table>
<thead>
<tr>
<th></th>
<th>All Intra Main</th>
<th>Random Access Main</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Y</td>
<td>U</td>
</tr>
<tr>
<td>Class A</td>
<td>-0.15%</td>
<td>-0.30%</td>
</tr>
<tr>
<td>Class B</td>
<td>-0.05%</td>
<td>0.10%</td>
</tr>
<tr>
<td>Class C</td>
<td>-0.14%</td>
<td>-0.54%</td>
</tr>
<tr>
<td>Class D</td>
<td>-0.07%</td>
<td>-0.58%</td>
</tr>
<tr>
<td>Class E</td>
<td>-0.20%</td>
<td>-1.22%</td>
</tr>
<tr>
<td>Class F</td>
<td>-1.56%</td>
<td>-1.11%</td>
</tr>
<tr>
<td>Overall</td>
<td>-0.36%</td>
<td>-0.55%</td>
</tr>
</tbody>
</table>

Table 3: Experimental results for LCU-based RD-optimization.

<table>
<thead>
<tr>
<th></th>
<th>All Intra Main</th>
<th>Random Access Main</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Y</td>
<td>U</td>
</tr>
<tr>
<td>Class A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Class B</td>
<td>-0.32%</td>
<td>-0.20%</td>
</tr>
<tr>
<td>Class C</td>
<td>-0.56%</td>
<td>-0.71%</td>
</tr>
<tr>
<td>Class D</td>
<td>-0.43%</td>
<td>-0.36%</td>
</tr>
<tr>
<td>Class E</td>
<td>-0.79%</td>
<td>-1.04%</td>
</tr>
<tr>
<td>Class F</td>
<td>-1.48%</td>
<td>-1.16%</td>
</tr>
<tr>
<td>Overall</td>
<td>-0.69%</td>
<td>-0.65%</td>
</tr>
</tbody>
</table>

Table 4: Experimental results for time consumption.

<table>
<thead>
<tr>
<th></th>
<th>Class A</th>
<th>Class B</th>
<th>Class C</th>
<th>Class D</th>
<th>Class E</th>
<th>Class F</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dec Time[%]</td>
<td>111.20%</td>
<td>115.67%</td>
<td>111.33%</td>
<td>111.20%</td>
<td>111.20%</td>
<td>111.20%</td>
<td>111.20%</td>
</tr>
<tr>
<td>Enc Time[%]</td>
<td>109.86%</td>
<td>112.85%</td>
<td>110.14%</td>
<td>110.14%</td>
<td>110.14%</td>
<td>110.14%</td>
<td>110.14%</td>
</tr>
</tbody>
</table>

Figure 8: Kimono, low-delay P, QP37, #32.

Figure 9: SlideEditing, random access, QP37, #40.