Gradient-based Offset Compensation for Edge Corrections on 8K Video Transmission using 4K Tiles

Hamid Reza Tohidypour, Yixiao Wang, Mahsa T. Pourazad, Chen Yuan, Runze Wang, Travis LePage, and Panos Nasiopoulos
Electrical & Computer Engineering, Univeristy of British Columbia
Vancouver, BC, Canada
{htohidyp,yixiaow,pourazad,cyuan}@ece.ubc.ca, runzw@student.ubc.ca, tlepage0105@gmail.com,panosn@ece.ubc.ca

***Abstract*— The increasing demand of emerging 8K video content, has made its transmission one of the main challenges for broadcasting companies. Thus, upgrading the existing infrastructures seem to be the only option to support transmission the 8K content at the desired quality level, which is costly and time consuming. To address this critical challenge, we propose a novel approach that avoids such a premature and costly upgrade of playback devices by splitting the 8K content into four 4K tiles which are encoded separately, transmitted, and merged at the decoder side. A unique gradient-based offset compensation technique is proposed to correct possible visual degradation at 4K tile boundaries. Our performance evaluation demonstrates the effectiveness of our approach under various 4K tile compression bitrates. This simple and yet efficient method also allows fast processing and real-time live broadcasting.**

Keywords*—— *8K video transmission; 4K tiles; deblocking; offset compensation component;

I. Introduction

With the continuous development of video and camera technologies in recent years, the video resolution has evolved from 1080p to the current 4K; 4K TVs and 4K set-top boxes have gradually become popular, and 4K content has gradually become abundant for normal uses. At the same time, 8K video technology has also been emerging, resulting in the appearance of 8K videos on mainstream platforms that provide online video services, such as YouTube and Netflix [1].

However, current broadcasting networks and infrastructures were designed and optimized for the transmission of 4K content. This means that they are not capable of supporting compressed 8K content at a pleasing quality level. This is the main challenge that the current broadcasting companies are facing while trying to adapt to the rapid pace of the multimedia technology with minimum cost.

In this paper, we propose a novel approach for transmitting 8K video, in which each 8K stream is divided into four 4K (equal sized) tiles which in turn are transmitted instead of the original 8K video. More precisely, each of these 4K tiles have to be compressed at the desired bit rate separately and merged later at the decoder side. The goal of this approach is to eliminate the costly need of upgrading existing playback devices for the near future. However, any tile based compression approach may result in full resolution frames with blocking artifacts at the edge (boundary) of the merged tiles. To address this issue, we proposed a low complexity gradient based offset compensation method to efficiently smooth the edges generated after merging the 4K tiles. Performance evaluations showed that our method outperformed the state-of-the-art deblocking approach, smoothing the edges at the boundary without degrading the visual quality of the 8K frame.

The rest of this paper is organized as follows. In Section II we give a detailed description of our proposed method. Section III presents the performance evaluation of our method and discusses the results. Finally, Section IV concludes our paper.

II. OUR PROPOSED APROACH

Our objective is to propose a method that allows video service providers to offer 8K content to their customers without going through the costly process of replacing the existing coding and transmission infrastructure and current video receivers. To this end, we propose to partition each frame of an 8K video into four 4K tiles. This results in four streams which will be separately encoded and transmitted to the end user, and then be reconstructed back to the original 8K resolution at the receiver end. This technique comes with its own set of issues, as compressing 4K tiles separately may result in artifacts and decorrelation in the reconstructed 8K video, especially along the edges of the separate tiles. This issue will be reviewed in this section, along with our recommended approach on how to resolve this matter. Although our proposed approach is tested for HEVC and high dynamic range (HDR), it can be easily used for other compression codecs and standard dynamic range (SDR) content. In the following subsections we first explain in detail the cause of possible blocking artifacts, and then the steps we took to design our proposed smoothing method.

A. Blocking artifact

Blocking artifacts are a result of the transformation and quantization stages of compression. During the HEVC compression process, individual video frames will be partitioned to blocks of pixels called coding blocks. These coding units will be further split into separate transform units. Each of the transform units will have the Discrete Cosine based Transform

979-8-3503-3959-8/23/$31.00 ©2023 IEEE

Fig. 1. 1MB/s compression on left vs 8MB/s compression on right.

applied to it, resulting in a matrix of transform coefficients [2]. The process of quantization during compression is the reason behind blocking artifacts seen in the compressed images. This process is done to each transform unit block separately without reference to adjacent blocks. This will cause decorrelation among the pixels of separate transform blocks, which will result in the transform blocks being visible in an image. This is especially problematic in regard to our 8K streaming technique. With each 4K tile going through a separate compression process, the edges of the tiles will be highly uncorrelated to adjacent tiles. As a result, there will be a clear "crosshair" in the 8K image where the four 4K tiles are coming together. An example of crosshair in the reconstructed 8K image is shown in Fig. 1 (left image). Note that the image is obtained by zooming in around the center area to make the artifact more visible.

B. *Effects of bitrate on the blocking artifact*

The first step in addressing this issue is to investigate how different compression bit rates can affect the edges in the 8K stream. Compressing at higher bit rates allows for more information to be retained from the original 4K videos, which will result in less of a perceivable edge along the center of the reconstructed 8K video. The tradeoff is more bandwidth will be required to transfer this information. This tradeoff will have to be considered throughout this paper, as the goal will be to find the minimum bit rate that our edge smoothing technique can still be successfully utilized. Fig. 1 shows the merged 4K tiles compressed at two different bitrates.

C. *Our proposed method*

1) Deblocking Using Offset and Shifting Techniques: One of the essential points in designing a deblocking approach for our application is that the approach should not mistakenly

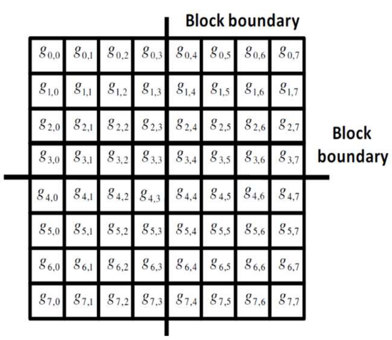


Fig. 2. Example showing pixel arrangement in a 8x8 section with blocking artifacts in the center.

TABLE I. CLASSIFICATION OF DEBLOCKING BLOCKS.

	UDB	HDB	VDB	CDB
ACTH	< T	> T	< T	> T
ACTV	< T	< T	> T	> T

consider real edges as the artificial ones generated after margining the tiles. In addition, the method should not be very complex and not introduce any significant delay at the receiver side. Our literature review showed that among all the existing deblocking methods, such as the in-loop and out-of-loop filters used in video compression standards [4][5], the method proposed by Kim et al. [3] could potentially match the requirements of our application.

Thus, we initially based our deblocking approach on this method. In order to perform deblocking using this method, first we define pixels near the block boundary as a deblocking block, which consists of 64 pixels: 8 pixel units in the horizontal direction and 8 in the vertical direction, as shown in Fig. 2.

The deblocking filters resulted in deblocking blocks alongside the block boundary. Each of the deblocking block is classified into four categories by calculating the differences between each pair of horizontally-adjacent/vertically-adjacent pixels and adding them together to get two values: horizontal activity (ACTH) and vertical activity (ACTV). Then, we compare these two values with a constant value T (threshold = 28) [3]. The approach to classify the deblocking blocks is shown in Table I. In this table, UDB, HDB, VDB and CDB represent uniform deblocking blocks, horizontal deblocking blocks, vertical deblocking blocks, and complex deblocking blocks, respectively.

For UDB and the directions with strong distortion (> T) in HDB and VDB, we apply a strong filter to remove the blocking artifact, the principle of which is shown in Fig.3. We select 8 nearest pixels to the block boundary. The idea is that the nearer a pair of pixels is to the boundary, the closer we move them together. For the direction with less distortion in UDB and VDB, we only modify the values of p3 and p4 in order not to change the real edges. For complex deblocking blocks (CDB), the new value of the pixel depends on all 8 adjacent pixels. The images before processing and after applying Kim's deblocking approach are shown in Fig. 4. As it can be seen, there is a clear

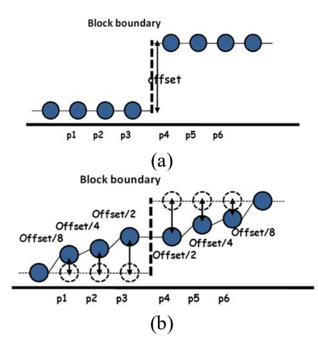
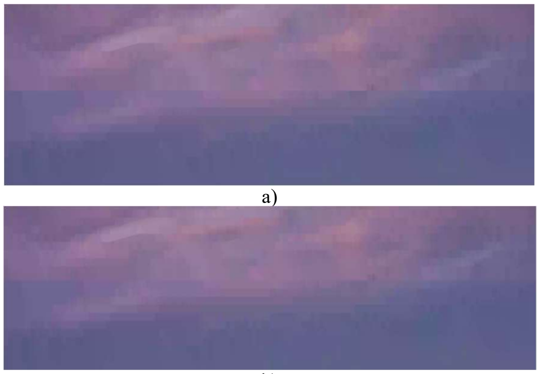


Fig. 3. a) pixel values at a boundary that introduces visual artifacts, b) shows the corrected pixel values by a strong filter.



a)

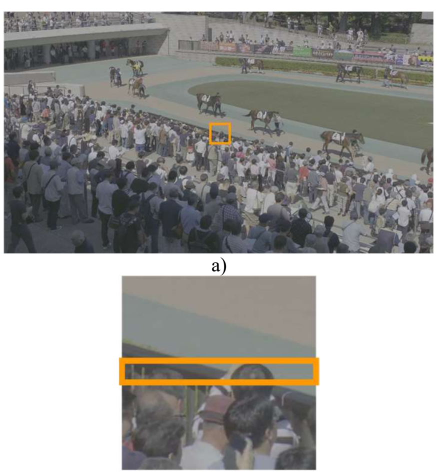
b)

Fig. 4. Image a) before processing, b) after processing.

horizontal edge in the middle of the image. After we apply offset and shifting techniques, the edge does become blurred, but it is still not too shallow to be invisible. This observation proves that this method can remove blocking artifacts, but the effect is not significant. So, in the next subsection we will talk about our improvement on this method.

2) Gradient-based Offset Compensation: As shown in Fig. 5, the edge in the center of the image can be reflected as the offset in the flow of the pixels in RGB domain. Fig. 6 shows the 2D plot of the pixel values across a horizontal line shown in Fig. 5 (of the zoomed area), where 0 represents the most left pixel of the zoomed area while 300 represents the most right pixel of the zoomed area. Since the offset in the Pixel RGB plot is corresponding to the edge in the image, from the pixel perspective, the elimination of the offset in the pixel plots can remove the edges in the center of the 8K videos. In addition, the pixel values should not be distorted to make sure the actual image information is maintained.

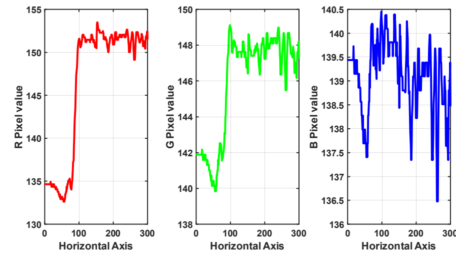
Across a horizontal center, the pixels on the left side of the plot, which is the top side of the image, are represented as



a)

b)

Fig. 5. a) An example four 4K tiles that are merged in to a 8K image and b) the zoomed area of edges. The 4K tiles were encoded at 12Mb/s.



a) R channel b) G channel c) B channel

Fig. 6. Corresponding pixel RGB plots across a horizontal line.

$q_1, q_2, \ldots, q_n$, while the pixels on the right side of the plot are represented as $p_1, p_2, \ldots, p_n$ (see Fig. 7 for more details). To remove the edges, the pixels on the left side, q_n, can remain unchanged, while the pixels on the right side, p_n, are updated based on the symmetrical pixels on the left side as well as the gradients of p_n and q_n which reflect the change of pixels on both sides.

The gradient-based offset compensation can be expressed in the RGB domain as follows:

$$p_1' = q_1 \quad (1)$$

$$p_n' = p_n + (q_n - p_n)/(1 + A_n), n > 1 \quad (2)$$

The updated pixels on the right side are represented by p_n'. If $n=1$, then (1) indicates that the first pixel on the right side should take the value of the first pixel on the left side. If $n>1$, the pixel on the right side is updated using the difference with the symmetrical pixel on the left side and a location-related term ($1+A_n$). An in the location related term is defined as follows:

$$A_n = \alpha \cdot \sum_n (|\Delta p_n| + |\Delta q_n|)^{\beta} + \alpha_1 \cdot (|\overline{p_n} - \overline{p_1}|)^{\beta_1} + \alpha_1 \cdot (|\overline{q_n} - \overline{q_1}|)^{\beta_1}, n > 1 \quad (3)$$

where Δp_n and Δq_n are the gradients of p_n and q_n, respectively. α, α1, β and β_1 are the constants to adjust the suppression, the values are 10^{-6},10^{-4}, 4 and 5, respectively. These values were empirically obtained using extensive subjective tests on a training dataset using involving 18 subjects. These values were obtained using empirical studies. ($\overline{p_n}$) and ($\overline{q_n}$) are the square root of sum of square of the RGB pixel values on the same location for p_n and q_n, as illustrated in the following equations:

$$\Delta p_n = p_n - p_{n-1} \quad (4)$$

$$\overline{p_n} = \sqrt{p_{n,R}^2 + p_{n,G}^2 + p_{n,B}^2} \quad (5)$$

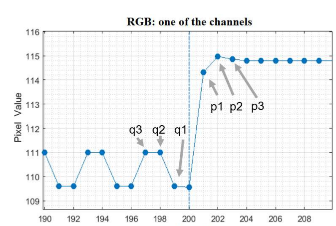


Fig. 7. The order of the pixels as they used by our proposed method.

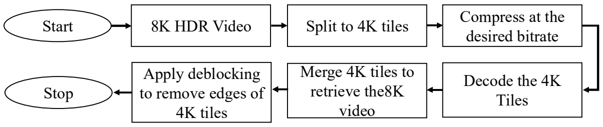


Fig. 8. Flowchart of our proposed approach.

The first item term of in (3) reflects the accumulated gradual change of pixel values results in the suppression of offset compensation while the second and third items terms reflect the rapid instant change results in more suppression of offset compensation. Fig. 8 shows the flowchart of our proposed approach for 8K video transmission.

III. PERFORMANCE EVALUATION AND DISCUSSION

For evaluating the performance of our deblocking method against the original merged 4K tiles and the Kim's deblocking approach [3], we decided to focus on subjective tests. This is the due to the fact that the ultimate users of our method will be human viewers. In addition, since the focus of our method is on the boundaries of the merged 4K tiles, which is a very small region compared to the entire 8K resolution, deblocking filters can slightly improve the quantitative results obtained by the objective metrics.

For our subjective test, we used a test dataset consists of six unseen 8K videos randomly selected from 8K video sequences of ITE [6] and the videos captured by the authors. Each of these videos were split into four 4K tiles and encoded at 6Mbs, 12Mbs and 20Mbs. The decoded tiles will be merged at the decoder side to form the 8K video. First, six 8K frames were randomly selected from each of these videos, resulting in 18 frames used in our subjective test. All of the 18 adult viewers that participated in our subjective test were screened for color blindness and visual acuity using the Ishihara and Snellen charts [7], respectively. Prior to conducting this experiment, subjects were trained to become familiar with the test procedure. The images before applying our deblocking method and after were shown on two identical Sony BVM-X300 professional monitors [8]. Subjects were first asked to compare the 8K frames obtained after merging the 4K tiles before and after applying our deblocking approach and choose one of the two images ("A" or "B") that looked more appealing to them in terms of overall quality or decide if they looked the same ("A=B"). Given the nature of the test, the subjects were instructed to focus mostly on the center of the image and the boundaries of the merged tiles. In the second test, subjects were asked to compare the 8K frames obtained after applying our deblocking approach and Kim's deblocking approach. Since these are 4K monitors, we center cropped each of the 8K images to extract 4K windows from the 8K frames. To validate the preference stability of the subjects and to consider any variations of brightness across the display, the position of the images was randomly changed between the right and the left monitors.

Next, we conducted subjective tests using the video sequences themselves. The process followed was identical to the one used for the frames, but in this case the subjects were asked to evaluate the quality of 10 seconds of randomly cropped section of each video. We used the outlier detection approach utilized in [7] to identify outliers. One outlier was detected for each of the tests and the corresponding results were excluded. Figs. 9 a) and b) show the average mean opinion score (MOS), representing the percentage of time subjects favored the visual quality of our method against the other approaches for the frame and video tests, respectively. We observe that, compared to the merging without deblocking, 88% of subjects favored our approach on images and 86.5% on videos on average. In addition, 84% and 82% of the subjects favored the visual quality of our deblocking approach against Kim's approach for the case of images and videos, respectively. It is worth mentioning that the proposed approach is characterized by minimal complexity. This is due to the fact that it solely involves horizontal and vertical activity classification based on the difference between adjacent pixels, as well as updating pixel values according to location and brightness difference. This makes it ideal for real-time applications.

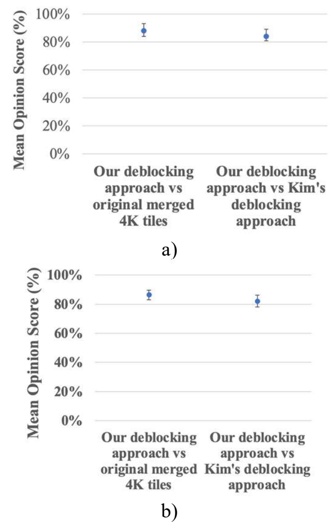


Fig. 9 Average mean opinion scores (MOS) comparing our method againgst merging 4K tiles without deblocking and Kim's deblocking approach for a) frames and b) videos.

IV. CONCLUSION

In this work, we proposed a unique approach for transmitting 8K video content via four 4K tiles to avoid the cost of upgrading encoding, decoding and transmission infrastructures. In our approach, the four 4K tiles are compressed separately and the decoded frames are merged into the original 8K resolution at the decoder side. We also introduce a novel gradient based offset compensation method that is shown to efficiently remove any possible blocking artifacts that may appear due to compression at the 4K tiles boundaries. Subjective performance evaluations showed that our proposed method works reliably under various compression bitrates and the compensation approach outperforms other state-of-the-art deblocking approaches in removing edge artifacts. The simplicity and computation effectiveness of our approach allows for small memory allocation and live broadcasting of the emerging 8K content.

References

[1] Y. Suzuki, H. Sujikai, “Transmission System of 4K/8K UHDTV Satellite Broadcasting.” IEICE Transactions on Communications, E103.B(10), 2020.

[2] G. J. Sullivan, J. -R. Ohm, W. -J. Han and T. Wiegand, "Overview of the High Efficiency Video Coding (HEVC) Standard," in IEEE Transactions on Circuits and Systems for Video Technology, vol. 22, no. 12, pp. 1649-1668, Dec. 2012.

[3] J. Kim, M. Choi and J. Jeong, "Reduction of Blocking Artifacts for HDTV using Offset-and-Shift Technique," in IEEE Transactions on Consumer Electronics, vol. 53, no. 4, pp. 1736-1743, Nov. 2007.

[4] G. J. Sullivan, J. -R. Ohm, W. -J. Han and T. Wiegand, "Overview of the High Efficiency Video Coding (HEVC) Standard," in IEEE Transactions on Circuits and Systems for Video Technology, vol. 22, no. 12, pp. 1649-1668, Dec. 2012.

[5] B. Bross et al., "Overview of the Versatile Video Coding (VVC) Standard and its Applications," in IEEE Transactions on Circuits and Systems for Video Technology, vol. 31, no. 10, pp. 3736-3764, Oct. 2021.

[6] ITE, Test chart Ultra-high definition, wide color gamut HDR version standard moving image (C series). https://www.ite.or.jp/content/test-materials/uhdtv_hdr/ (accessed June. 2, 2022).

[7] P. Mohammadi, M. T. Pourazad and P. Nasiopoulos, "A Fully Automatic Content Adaptive Inverse Tone Mapping Operator With Improved Color Accuracy," in IEEE Transactions on Circuits and Systems for Video Technology, vol. 31, no. 11, pp. 4198-4213, Nov. 2021

[8] SONY BVM-X300 4K OLED Master Monitor. [Online]. Available: https://pro.sony/. Accessed: December 2021.