INTERNATIONAL ORGANISATION FOR STANDARDISATION
ORGANISATION INTERNATIONALE DE NORMALISATION
ISO/IEC JTC1/SC29/WG11
CODING OF MOVING PICTURES AND AUDIO

ISO/IEC JTC1/SC29/WG11 MPEG2019/m50818

October 2019, Geneva, CH

Source **Sungkyunkwan University(SKKU)**

Title **CE3: Viewport-dependent Patch Grouping Using HEVC Tiles**

Authors **Soonbin Lee, Jong-Beom Jeong, Eun-Seok Ryu**

1 Patch-Group Tile (PGT) Method

After 3DoF+ call for response, the *Group-Based TMIV* [1, 2] has been proposed to MPEG-Immersive standard meetings by Intel Corporation. The group-based TMIV method preprocesses view group information separately and merges the metadata of these view groups that leads the subjective video quality improvement. However, the group-based TMIV method has the limitation that needs multiple encoders and decoders because of multi-pass encoding approach.
This proposal introduces the viewport-dependent patch grouping method using HEVC/VVC Tiles. Compared to the group-based TMIV method, the proposed method packs patches into Tiles independently considering each view Tile. Thus, in viewport-dependent virtual reality(VR) streaming systems, the server does not need to transmit all Atlas patches with metadata but just needs one motion-constrained tile set (MCTS) encoder and decoder that are already applied to HEVC test model (HM).
The user viewport (views) can be classified as groups, and the proposed method packs the patches into Tiles for each group. Each Tile(s) can be transmitted to the client head-mounted display (HMD) independently using the Tile extractor. Figure 1 shows the proposed encoder architecture that uses single encoder and decoder.

Patch-Group Tile (PGT) stands for the HEVC/VVC Tile(s) filled with patch group(s). Mostly, patches can be grouped according to views. Here are advantages of the proposed PGT method.

Advantages:

- Single-pass encoding/rendering with MCTS-encoder and modified TMIV.
- Significantly improved subjective quality.
- Significant bandwidth saving with selective PGT streaming.

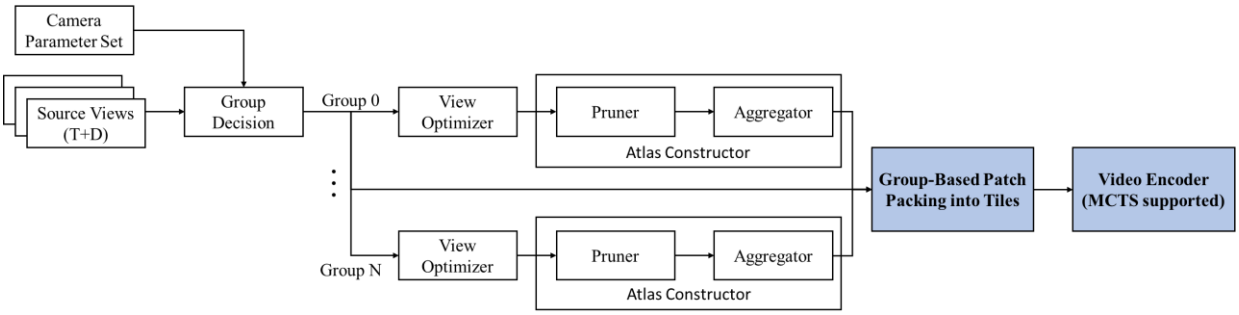

Fig. 1. Encoder architecture of the prposed viewport-dependent patch grouping method using HEVC/VVC Tiles.

1.1 Tiled basic views and PGT method with Atlas patch groups packed into Tiles

Based on Basel et al.[1], 3DoF+ test materials are divided into groups, and the group-based TMIV increased the subjective quality.

The PGT method enhances the concept of the group-based approach with (1) Tiled basic views and (2) Atlas patch groups packed into Tiles. The method requires the motion-constrained Tile set (MCTS)-enabled video encoder and provides significant bandwidth savings when selected Tile(s) is(are) transmitted to the HMD according to user viewport. In addition, the proposed method does not need multi-pass encoding as [1] but single-pass encoding.

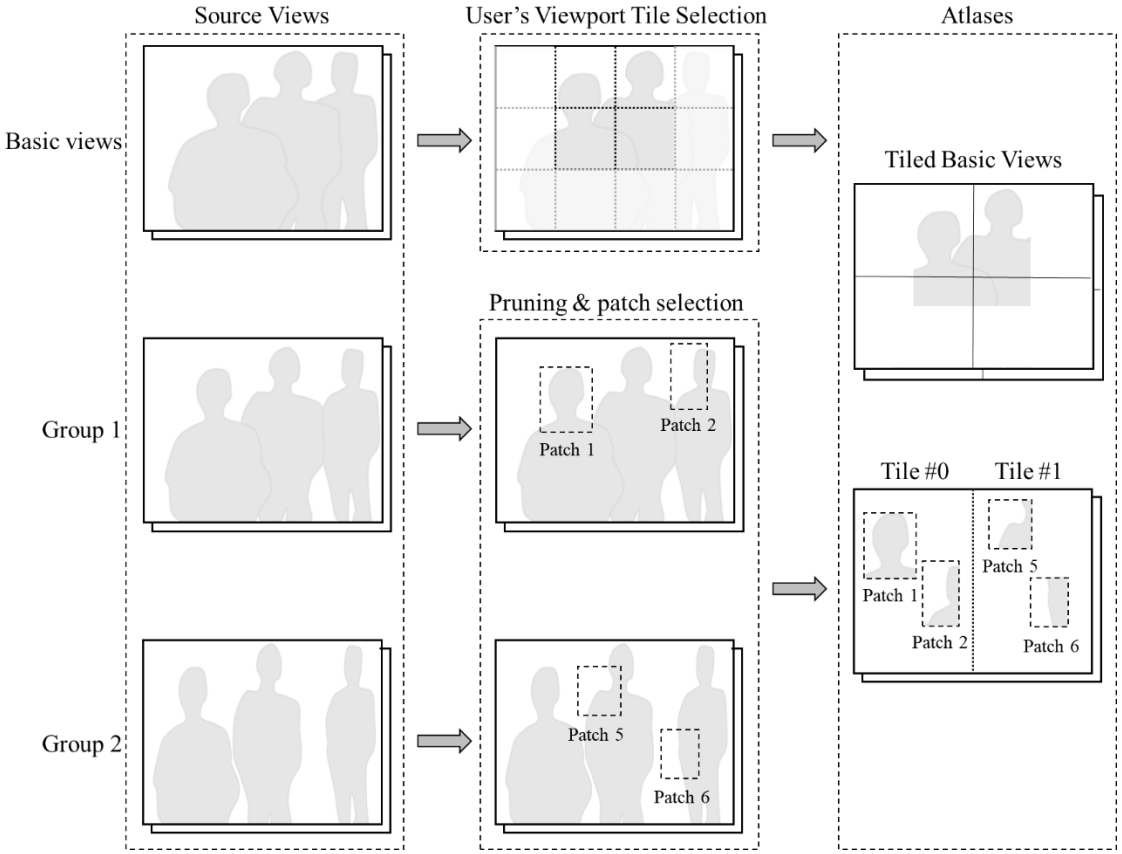

Fig. 2. Conceptual diagram of Tiled basic views and PGT method with Atlas patch groups packed into Tiles.

Modified TMIV v1.0 and HEVC reference SW with MCTS (HM 16.20) are used in our experiments, and the SW and results will be provided to next 129th MPEG meeting. The HM SW also supports Tile bitstream extractor, and each extracted Tile can be transmitted to client. Though the bitrate of whole Atlas is higher than method [1] because of the limited motion prediction of MCTS between Tiles, the actual required bandwidth can be much lesse than [1] because the proposed method transmits viewport Tile(s) only.
Figure 2 shows the conceptual diagram of Tiled basic views and PGT method with Atlas patch groups packed into Tiles. In the figure, Basic views and patches are divided into multiple Tiles separately.

2 Proposed Syntax and Semantics

The proposed syntax is as shown in Table 2 and 3. The *uniform_flag* indicates whether Tiles are divided uniformly or not. The number of Tiles can be more than one. Tile needs its own width and height (multiple of CTU (64 pixels)). Each patch knows the Tile identification (ID) that it belongs.

Table 2. Atlas Parameters List Syntax

atlas_params_list( ) {	
num_atlases_minus1	ue(v)
~~**num_groups**~~	~~u(8)~~
for (i = 0; i <= num_atlases_minus1; i++) {	
atlas_id[i]	ue(v)
~~**group_id**[i]~~	~~u(v)~~
uniform_flag[i]	u(v)
num_tile_columns_minus1[i]	u(v)
num_tile_rows_minus1[i]	u(v)
for (j = 0; j <= num_tiles_minus1; j++) {	
tile_id[i][j]	u(v)
tile_width[i][j]	u(v)
tile_height[i][j]	u(v)
tile_pos_in_atlas_x[i][j]	u(v)
tile_pos_in_atlas_y[i][j]	u(v)
}	
atlas_params(atlas_id[i])	
}	
}	

uniform_flag[i] specifies the tile sizes are uniformly distributed in i-th atlas.
num_tiles_columns_minus1[i] specifies the number of columns in i-th atlas.
num_tiles_rows_minus1[i] specifies the number of rows in i-th atlas.
tile_id[i][j] specifies the tile index of the j-th tile in i-th atlas.
tile_width[i][j] indicates the width of the j-th tile in i-th atlas.
tile_height[i][j] indicates the height of the j-th tile in i-th atlas.

tile_pos_in_atlas_x[i][j] and **tile_pos_in_atlas_y[i][j]** specify the horizontal and vertical coordinates in luma samples, respectively, of the top-left corner of the j-th pixel of the i-th atlas.

3 Conclusion

This proposal introduces the Patch-Group Tile (PGT) method using HEVC/VVC Tile(s) filled with patch group(s). Mostly, patches can be grouped according to views.
The proposed method has advantages: (1) provides single-pass rendering with modified TMIV, (2) improves subjective quality significantly, and (3) saves bandwidth significantly with selective PGT streaming according to user viewport.

References

[1] B. Salahieh, S. Bhatia, J. Boyce, "Group-Based TMIV," ISO/IEC JTC1/SC29/WG11 MPEG/m49406, July 2019, Gothenburg, Sweden.

[2] B. Salahieh, S. Bhatia, J. Boyce, "Grouping Implementation in TMIV," ISO/IEC JTC1/SC29/WG11 MPEG/m49859, July 2019, Gothenburg, Sweden.

[3] B. Salahieh, B. Kroon, J. Jung, and M. Domański, "Test Model for Immersive Video," ISO/IEC JTC1/SC29/WG11 MPEG/N18470, April 2019, Geneva, Switzerland.

[4] J. Jung, B. Kroon, and J. Boyce, "Common Test Condition for Immersive Video," ISO/IEC JTC1/SC29/WG11 MPEG/N18443, April 2019, Geneva, Switzerland.

[5] J. Boyce, R. Doré, and V. K. M. Vadakital, "Working Draft 2 of Metadata for Immersive Media (Video)", ISO/IEC JTC1/SC29/WG11 MPEG/N18576, July 2019, Gothenburg, SE.

[6] "Requirements for Immersive Media Access and Delivery", July 2019, N18654, Gothenburg, SE.