

A New Video Coding Scheme using Warped Reference Pictures [JVET-J0046]

Jung Won Kang (ETRI)

Dae Young Lee, Tea Hyun Kim, Gwang Hoon Park (KHU)



Contents

- ☐ Introduction
- ☐ Problem Statement
- ☐ Warped Picture for Reference
- ☐ Proposed Coding Scheme
- ☐ Experimental Results
- ☐ Further Study

Introduction

- ❑ Trend-shift of captured video sequences by users using from tripod-based to nontripod-based cameras
 - ❖ Nontripod-based cameras are easier to move and more likely to create complex movement than tripod-based cameras



Tripod-based cameras



Nontripod-based cameras
(Smartphone, drone, action cam)

- ❑ Various videos with global motion due to camera movement (example)



Tripod-based capture



Nontripod-based captures
by Dron



Nontripod-based captures
by Action Cam

Problem Statement

❑ Global motion by camera movement

❖ Camera movement causes global motion

- With nontripod-based cameras, camera movements occur very often
- Tripod-based cameras can also cause camera movement by camera work

❖ Type of global motion

- Linear global motion : panning, tracking
- Nonlinear global motion : rotation, dolly-in(zoom-in), dolly-out(zoom-out)
- In the conventional method, it is difficult to obtain a high coding efficiency when nonlinear global motion occurs



No global motion



Linear global motion



Nonlinear global motion
dolly-in by car black box



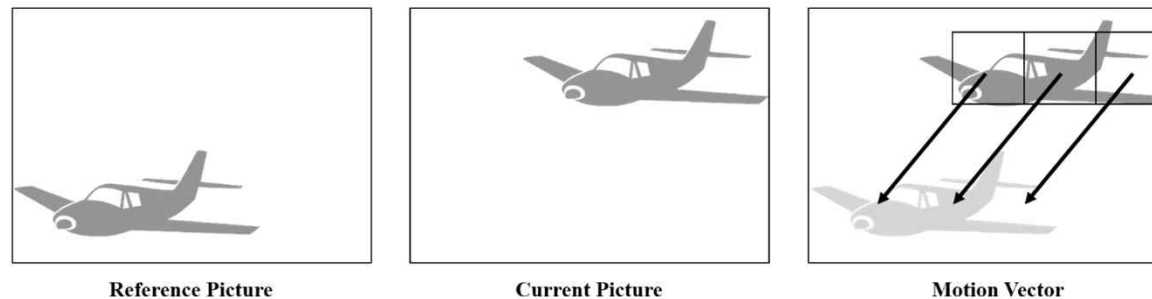
Nonlinear global motion
rotation by dron

Problem Statement

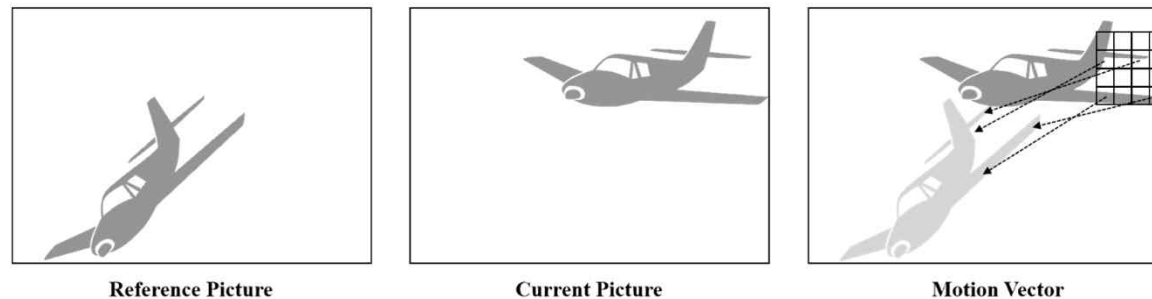
❑ Inter prediction when nonlinear motion occurs in HEVC

❖ Motion Vector derivation examples in HEVC

- The prediction region is more finely subdivided when nonlinear motion occurs



(a) Motion vector derivation when linear motion occurs

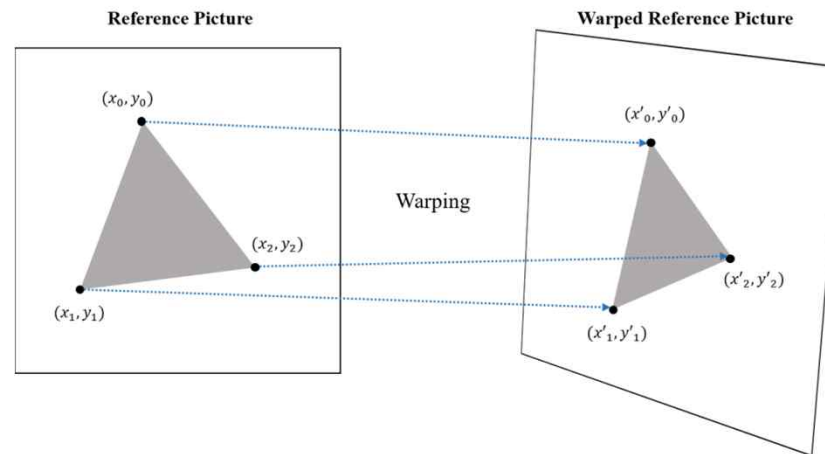


(b) Motion vector derivation when non-linear motion occurs

Warped Picture for Reference

❑ Warped reference picture (WRP)

- ❖ Warped reference picture made from reference picture for better prediction
- ❖ Reflects geometric movements
 - rotation, translation, enlargement, reduction, projection
- ❖ As a model for expressing geometric transformation relations, homography is used
 - Geometric transformation relations are represented by 3x3 matrix
 - matrix generated by homography (warping matrix)

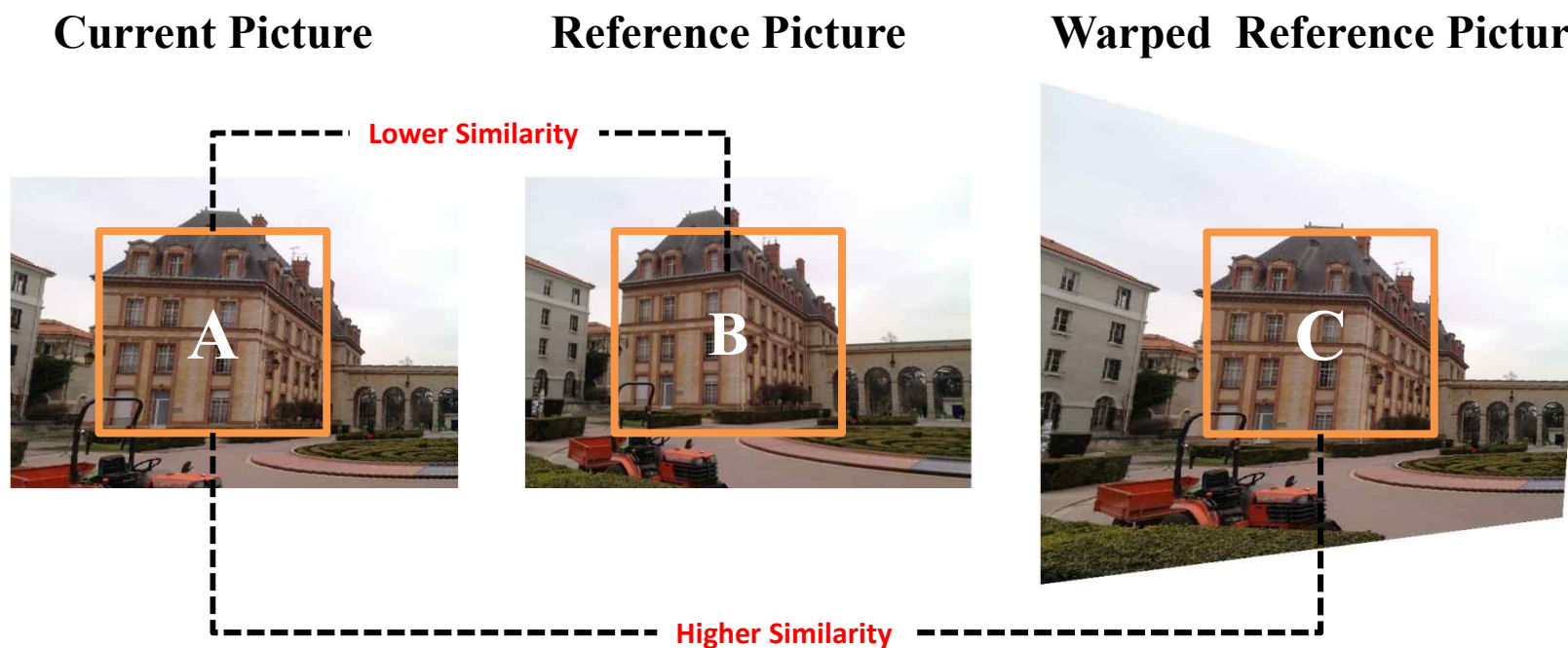


$$\begin{aligned} x'_n &= f(x_n, y_n, \alpha) \\ y'_n &= g(x_n, y_n, \beta) \end{aligned} \rightarrow \begin{bmatrix} x'_n \\ y'_n \\ 1 \end{bmatrix} = H \begin{bmatrix} x_n \\ y_n \\ 1 \end{bmatrix} \quad H = \begin{bmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ h_7 & h_8 & h_9 \end{bmatrix}$$

Warped Picture for Reference

❑ Features for WRP

- ❖ Higher similarity to the current picture compared to reference picture
- ❖ Reflects global movement
 - Linear movement : panning, translation
 - Nonlinear movement : rotation, zoom-in, zoom-out, dally-in, dally-out, projection



Proposed Coding Scheme

❑ Reference picture structure with WRP

❖ WRPs are generated from reference pictures using warping matrix

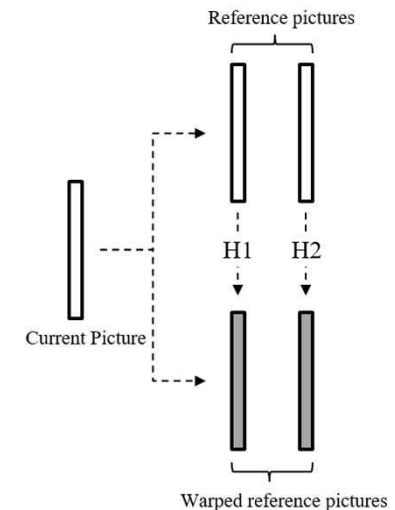
- Warping matrix is signalled in slice header
- Each warping matrix takes 256 bits (32bits/parameter)
- Each slice has maximum of 4 warping matrices (1,024 bits)

$$H = \begin{bmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ h_7 & h_8 & 1 \end{bmatrix}$$

❖ Current picture can be predicted from both reference picture and WRP

❖ Warping matrix is only signalled when WRPs are used for inter prediction in the picture

❖ Warping process of generating WRP is performed once if there is the same reference picture in L0 and L1

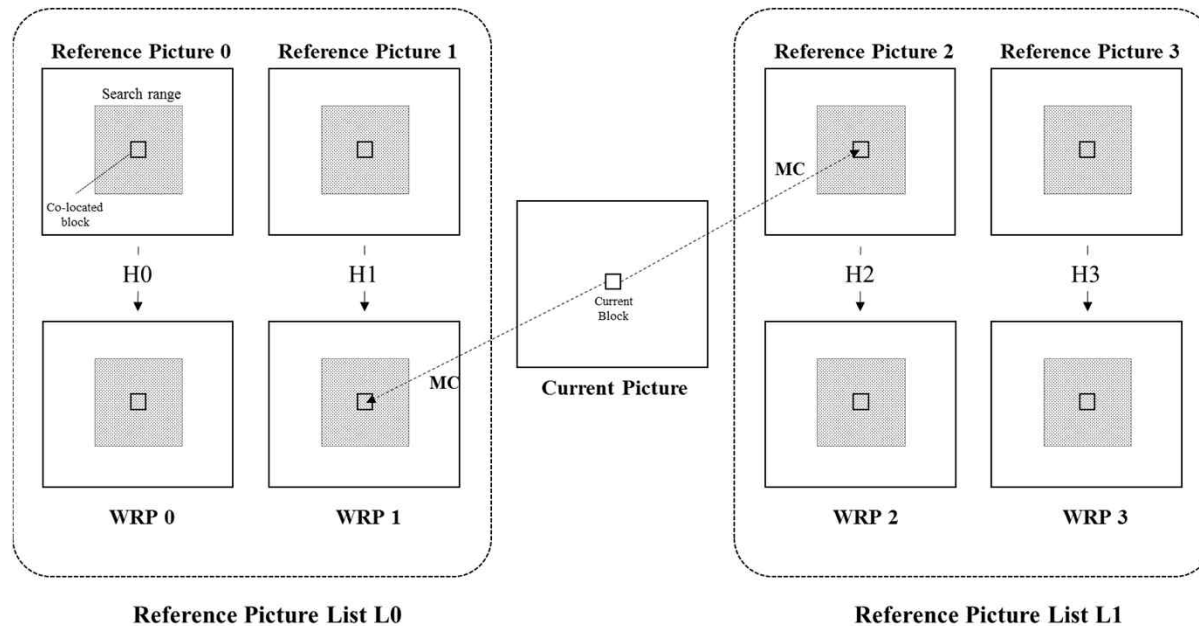


Proposed Coding Scheme

❑ Inter predictions using WRP

❖ Motion Compensation based on WRPs and general reference pictures

- WRP flag is signalled to indicate whether WRP is used for a PU
 - When WRP flag is true, WRP is used for the PU
 - When WRP flag is false, general reference picture is used for the PU
- Reference list number(L0, L1) and reference picture index of the WRP are the same as the corresponding reference picture



Proposed Coding Scheme

❑ Inter predictions using WRP

❖ Merge with WRP selection

- When a current PU is coded as merge mode, WRP flag is also copied along with motion information from the corresponding merge candidate

Experimental Results

□ Performance

❖ Sequence groups

Test Groups	Video Sequences	Remarks
Group A	JVET CTC test sequences	
Group B	JVET sequences with global motion	
Group C	KHU-created video sequences with global motion	Can be donated to JVET

❖ BD rate(Y) compared to HM16.9

- Group A: -0.68%
- Group B: -11.54%
- Group C: -22.88%

❖ Group A has mostly no global motion

- Low coding efficiency improvement for sequences without global motion

❖ Group B & Group C contain various degrees of global motion

- Higher coding efficiency compared to Group A

Experimental Results

❑ Group A : JVET CTC test sequences

Group A	BD-rate (%) for RA (Anchor : HM16.9)			EncT (%)	DecT (%)
	Y	Cb	Cr		
Class A1	-0.03	-0.04	-0.06	321	161
Class A2	-3.22	-3.19	-3.11	320	174
Class B	-0.02	-0.41	-0.32	359	138
Class C	-0.25	-0.71	-0.81	332	137
Class D	-0.56	-0.85	-1.17	345	142
Class E	0.03	0.05	0.01	243	138
Overall	-0.68	-0.88	-0.82	325	148

❑ Group C : KHU-created sequences

Group C	BD-rate (%) for RA (Anchor : HM16.9)			EncT (%)	DecT (%)
	Y	U	V		
KHU_Rotate	-32.83	-27.36	-27.54	232	364
KHU_Zoom_in	-16.91	-13.11	-13.36	230	260
KHU_Zoom_out	-15.59	-12.97	-13.32	226	296
Overall	-22.88	-18.77	-19.02	229	313

❑ Group B : JVET sequences with global motion

Group B	BD-rate (%) for RA (Anchor : HM16.9)			EncT (%)	DecT (%)
	Y	U	V		
BlueSkyPart	-16.22	-17.31	-13.94	196	468
JetsPart	-7.56	-8.71	-6.32	134	177
ShieldsPart	-18.99	-12.06	-10.86	189	401
SpincalendarPart	-15.87	-18.52	-21.76	183	230
StationKtaPart	-31.71	-25.30	-28.69	200	483
TractorPart	-18.39	-16.37	-16.28	232	347
IceAerial	-13.28	-11.51	-10.99	216	518
IceRiver	-4.75	-5.49	-5.41	227	263
IceRock2	-3.46	-3.67	-3.38	218	236
IceRock	-1.82	-2.06	-1.99	212	161
Netflix_Aerial	-15.65	-14.79	-14.46	197	326
Netflix_FoodMarket2	-3.55	-4.74	-4.26	230	366
Netflix_FoodMarket	-9.62	-6.74	-7.84	239	459
Netflix_RitualDance	-0.76	-0.73	-0.49	280	163
Overall	-11.54	-10.57	-10.48	211	328

Experimental Results

❑ Visualized analysis

❖ MV information

- Since WRP already reflects global motion, each MV becomes small and more similar to the neighboring PU
- Number of MVs generated by PU segmentation is decreased

❖ Merge region is in a wider area

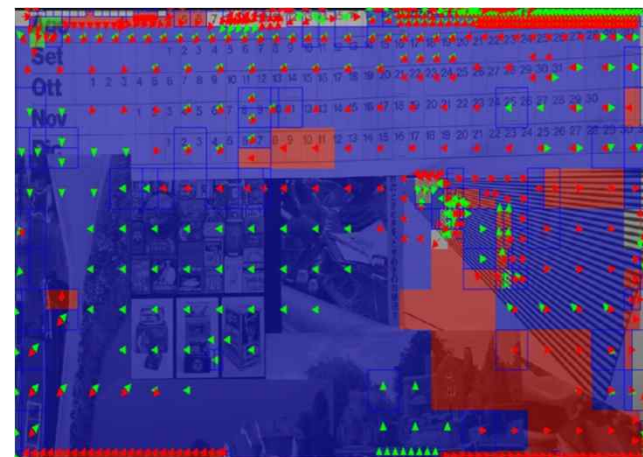
Experimental Results

❑ Visualized analysis - SpincalendarPart POC 4 (Partial Rotation)

HM 16.9

Proposed

Motion Vector



Merge Group



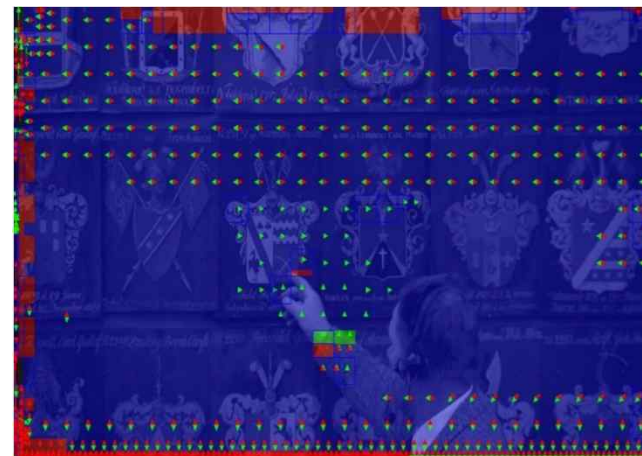
Experimental Results

❑ Visualized analysis - ShieldsPart POC 50 (Partial Zoom-in)

HM 16.9

Proposed

Motion Vector



Merge Group



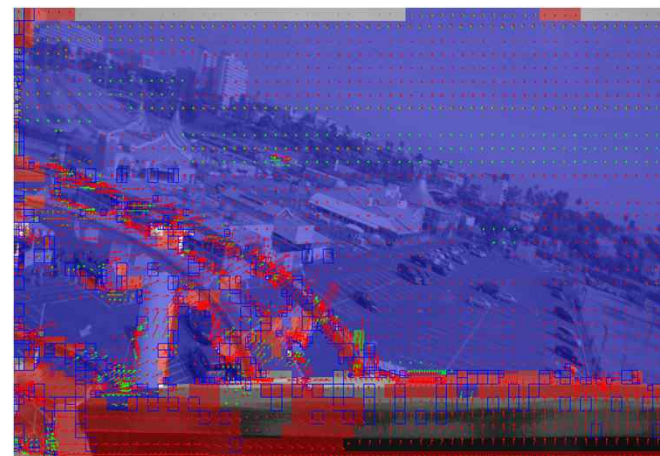
Experimental Results

❑ Visualized analysis - RollerCoaster POC 29 (Partial Projection)

HM 16.9

Proposed

Motion Vector



Merge Group



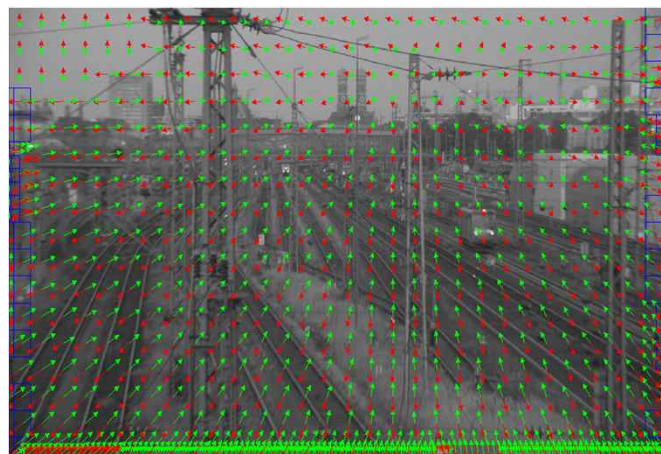
Experimental Results

❑ Visualized analysis - StationKtaPart POC 1 (Partial Zoom-out)

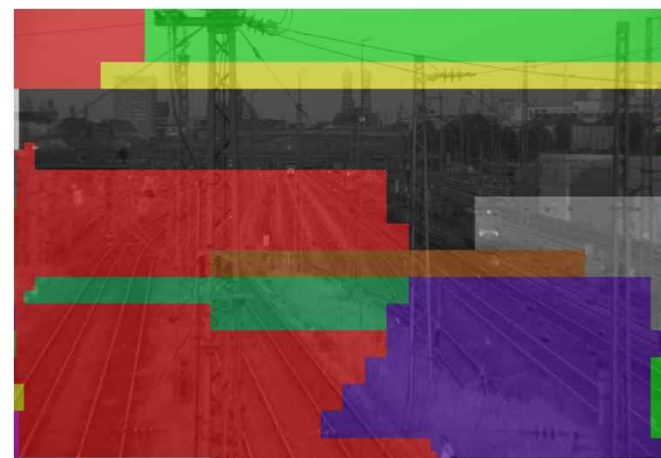
HM 16.9

Proposed

Motion Vector



Merge Group



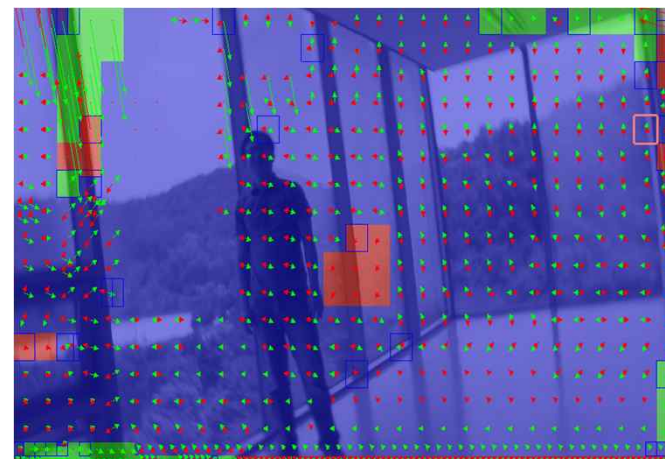
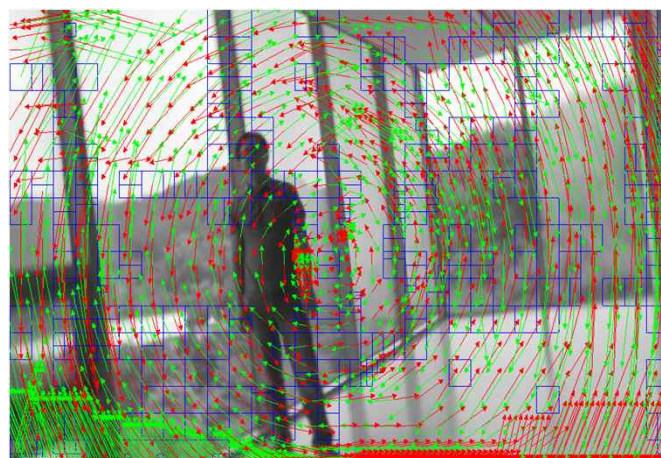
Experimental Results

❑ Visualized analysis – KHU_Rotate_fast POC 19 (Fast Rotation)

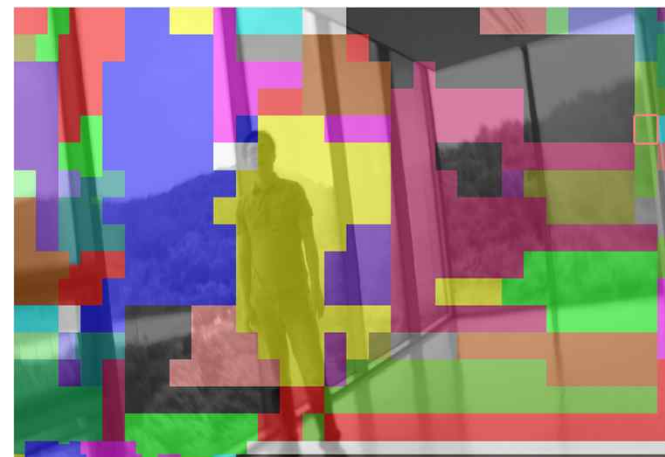
HM 16.9

Proposed

Motion Vector



Merge Group



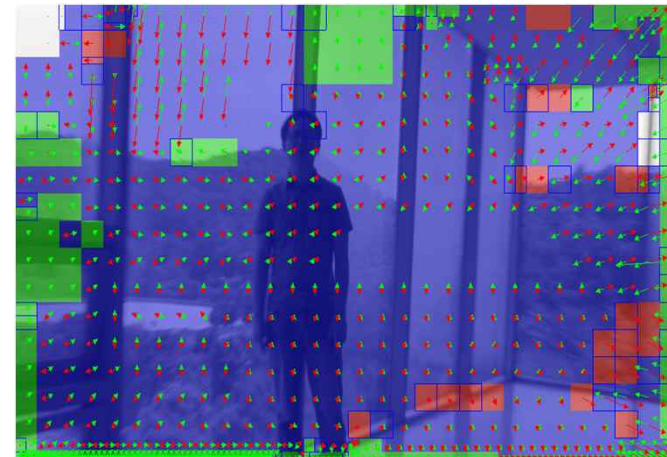
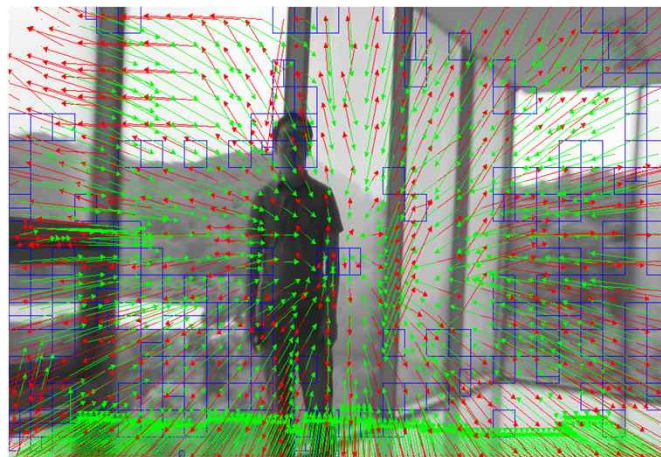
Experimental Results

❑ Visualized analysis – KHU_Zoom_out_fast POC 1 (Fast Zoom-out)

HM 16.9

Proposed

Motion Vector



Merge Group



Further Study

❑ Issues

❖ Extend block size

- Proposed method tends to minimize PU segmentation in video sequences with nonlinear global motion
- Coding efficiency can be improved with larger CTU size
 - In the proposed method, CTU size is 64x64

❖ Parallel processing

- The process for acquiring WRP at encoder can be parallelized in task-level
 - It is frame-based operation for each reference frame
- ❖ At decoder, WRP can be generated on block basis instead of frame basis
 - Only for a block using WRP as a reference picture, the Warped Reference Block can be generated

Further Study

❑ Issues

❖ Appropriate entropy coding

- Need entropy coding suitable for the characteristics of the proposed method

❖ Warping matrix compression

- Need efficient coding for warping matrix
 - In the proposed method, warping matrix is not compressed
 - Each warping matrix requires 256 bits (32 bits/parameter) and each slice has maximum of 4 warping matrix (1,024 bits)
- Bits for the warping matrix are considered to decide whether to use WRP.
If effective compression for warping matrix parameters is applied, the WRP will be applied more.

	QP	Ratio of bits for warping matrix over total bits (%)
Group A	22	0.15
	27	0.16
	32	0.32
	37	0.51
	Total	0.29

	QP	Ratio of bits for warping matrix over total bits (%)
Group B	22	0.82
	27	1.68
	32	2.87
	37	4.39
	Total	2.44

	QP	Ratio of bits for warping matrix over total bits (%)
Group C	22	1.27
	27	2.55
	32	4.58
	37	7.06
	Total	3.86

