MPEG-Immersive 6DoF Standard Work: Related to MPEG Immersive Video

Presenter: Eun-Seok Ryu (esryu@skku.edu)

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

Multimedia Computing Systems Lab. (MCSL)

http://mcsl.skku.edu

Department of Computer Education

Sungkyunkwan University

Compliance with IEEE Standards Policies and Procedures

Subclause 5.2.1 of the *IEEE-SA Standards Board Bylaws* states, "While participating in IEEE standards development activities, all participants...shall act in accordance with all applicable laws (nation-based and international), the IEEE Code of Ethics, and with IEEE Standards policies and procedures."

The contributor acknowledges and accepts that this contribution is subject to

- The IEEE Standards copyright policy as stated in the *IEEE-SA Standards Board Bylaws*, section 7, http://standards.ieee.org/develop/policies/bylaws/sect6-7.html#7, and the *IEEE-SA Standards Board Operations Manual*, section 6.1, http://standards.ieee.org/develop/policies/opman/sect6.html
- The IEEE Standards patent policy as stated in the *IEEE-SA Standards Board Bylaws*, section 6, http://standards.ieee.org/guides/bylaws/sect6-7.html#6, and the *IEEE-SA Standards Board Operations Manual*, section 6.3, http://standards.ieee.org/develop/policies/opman/sect6.html

MPEG-Immersive 6DoF Standard Work: Related to MPEG Immersive Video

Date: 2021-07-19

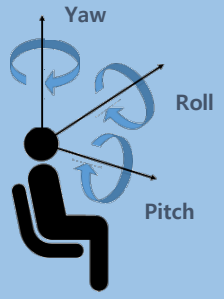
Author(s): Jong-Beom Jeong, Soonbin Lee, Eun-Seok Ryu

Name	Affiliation	Phone [optional]	Email [optional]
Jong-Beom Jeong	Sungkyunkwan Univ.		uof4949@skku.edu
Soonbin Lee	Sungkyunkwan Univ.		soonbinlee@skku.edu
Eun-Seok Ryu	Sungkyunkwan Univ.		esryu@skku.edu

MPEG Immersive Video (MIV) Overview

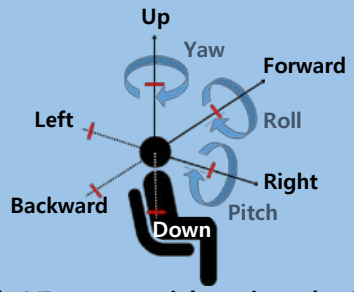
- Step-by-step objective of ISO/IEC MPEG immersive video
  - MPEG-I is responsible for standardizing immersive media in MPEG and specifies the goals of step 3
  - Goal of Revitalizing VR Commercial Service by 2020
  - Goal of 6DoF media support by 2022 after completing 3DoF standard by 2017

Step 1



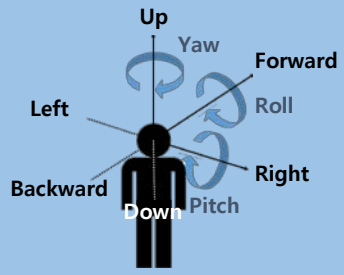
- Complete 3DoF standard by 2017
- Rotate head in a fixed state
- 360 video full streaming by default
- Tiled streaming if possible

Step 2



- Enable VR commercial servicers by 2020
- Allow head rotation and movement within a restricted area
- User-to-user conversations and projection optimization

Step 3



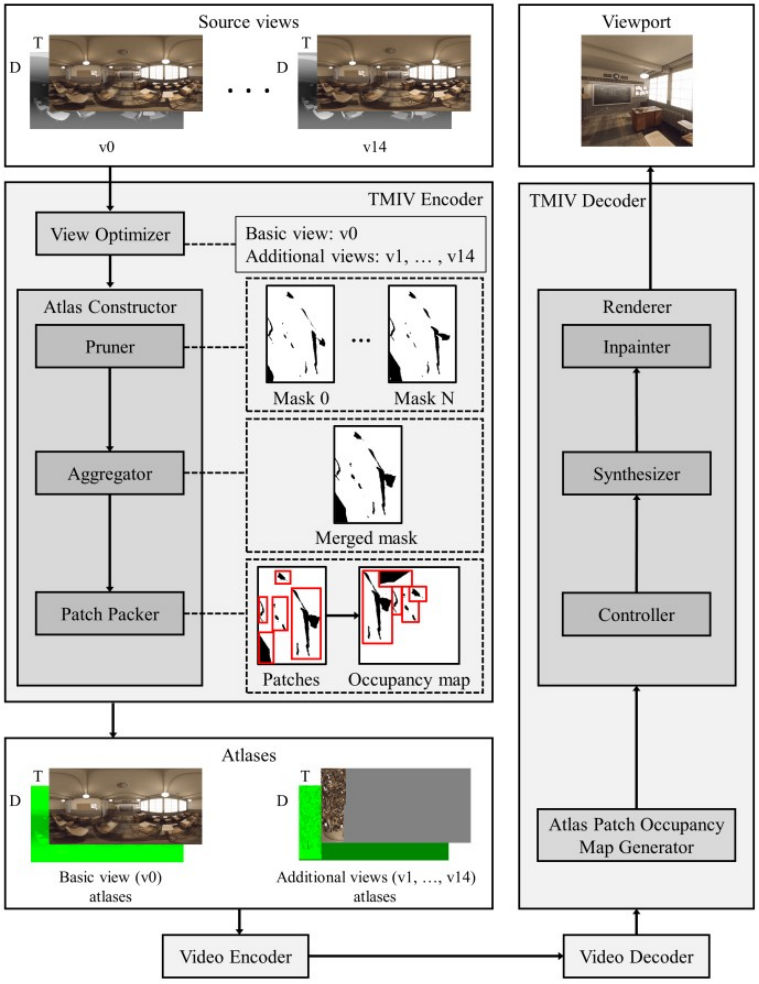
- Support 6DoF by 2022
- 6DoF video will reflect user's walking motion
- Support interaction with virtual environments

Definitions of 3DoF, 3DoF+, and 6DoF

MPEG Immersive Video (MIV) Overview (-Cont'd)

- MPEG-I requested call for proposals (CfP) on 3DoF+
- Philips, Intel & Technicolor, Nokia, PUT & ETRI, ZJU submitted responses > test model for immersive video (TMIV)
- Removes correlation between multi-view
- VR streaming on 6DoF natural-content

Viewport generated by TMIV



Block diagram of TMIV architecture

MPEG Immersive Video (MIV)-related Activities (-Cont'd)

- Mandates and Input Contributions in MPEG134 MPEG-I Visual
  - 1. Prepare draft international standard of MPEG Immersive Video 2
  - 2. Study profiles, tiers, and levels for MPEG Immersive Video 1
  - 3. Carry on core experiments of MPEG Immersive Video 3
  - 4. Update test model and reference software for MPEG Immersive Video 3
  - 5. Define CTC for MPEG Immersive Video 8
  - 6. Carry out exploration experiments on MPEG Immersive Video 10
  - 7. Investigate dense light field content in immersive video applications 10
  - 8. Study performance improvements under CTC conditions 2
  - 9. Promote MPEG Immersive Video outside MPEG 0
  - Total 39 doc.

MIV Common Test Conditions

- Reference softwares

Software name	Location	Tag/branch
TMIV	https://gitlab.com/mpeg-i-visual/tmiv	v8.0
VVenC	https://github.com/fraunhoferhhi/vvenc	v0.2.0.0
VVdeC	https://github.com/fraunhoferhhi/vvdec	v0.2.0.0
VMAF	https://github.com/Netflix/vmaf	v1.3.14
IV-PSNR	https://gitlab.com/mpeg-i-visual/ivpsnr	v3.0
IVDE	https://gitlab.com/mpeg-i-visual/ivde	v3.0

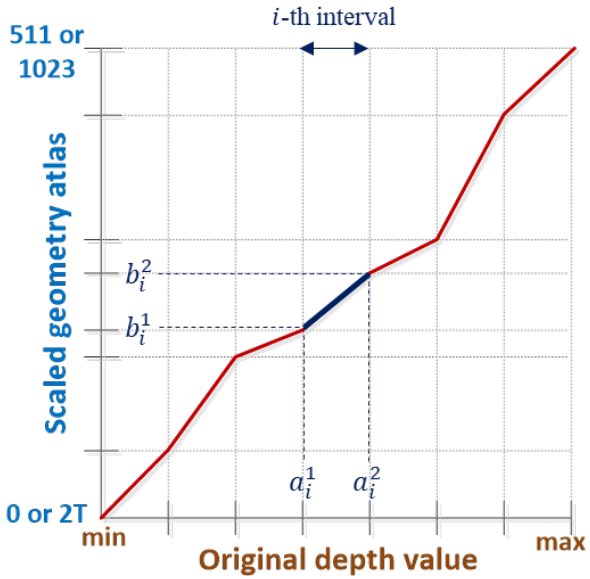
[MIV] ZJU response to CE2 viewport culling (m56415)

- Previous culler removed some patches from the viewport
- This is because of the patches which has FoV >= 180°
- Divided such large patches into sub-blocks
- AhG recommendations:
  - Integrate into TMIV
  - Enable in the CTC anchors

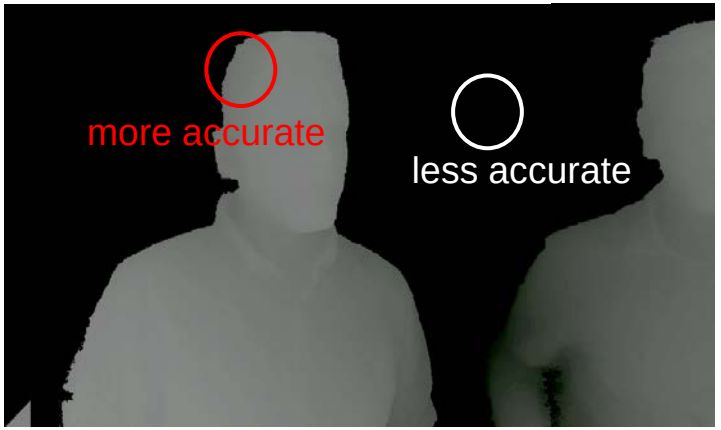
Frame in SQ P01 anchor (left) and modified subblock culler (right)

[MIV] CE3.2 Piecewise Linear Scaling of Geometry Atlas (m56481)

- Current TMIV normalizes depth values into 10-bits
- Proposed method conducts:
  - Divide depth into predefined number of intervals (proposed: 16)
  - Adaptively scale depth values
- Important close depth - disparity is large



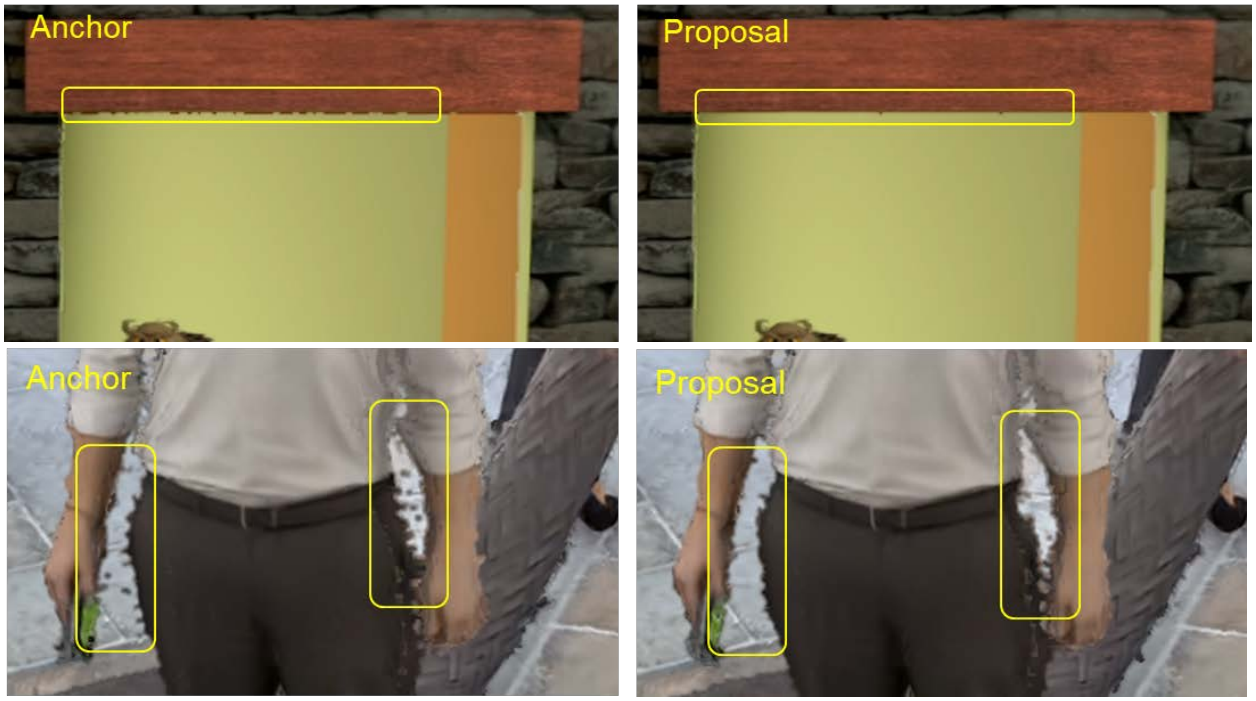
$$d' = \frac{(b_i^2 - b_i^1) * (d - a_i^1)}{a_i^2 - a_i^1} + b1_i$$



Example of piecewise linear scaling

[MIV] CE3.2 Piecewise Linear Scaling of Geometry Atlas (m56481) (-Cont'd)

- 3.4% BD-rate gain compared to the anchor
- Subjective quality advantages at the edges
- Proposed new syntax for depth quantization



Left: anchor viewport, right: proposed viewport

Pose trace issues with TMIV8 (m56774)

- Block artifacts are generated when using TMIV8 anchor (VVenC)
  Maybe because of optimization in VVenC..? Blocky artifacts in high QP
- AhG recommendations:
  - 1. Better crosschecking:
    Evaluate the pose traces, study the code, apply a strict expert viewing process
  - 2. Focus more on the pose trace generation:
    Viewing pose traces when releasing TMIV9, objective result is less important
  - 3. Tool evaluation:
    Decision with only subjective results, encode 17 frames and render 97

Frog pose trace P01 in QP5

Frame Packing Implementation in TMIV (m56827)

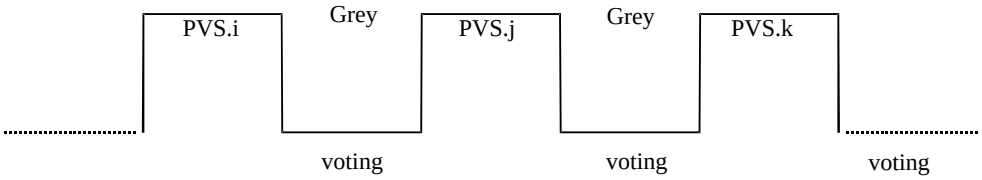
- Frame packing implementation in TMIV to:
- No empty spaces in the packed video
- Not restricted to VVC (sub-pictures) or HEVC-specific (MCTS) solutions
- Frame packing across different atlas indices is not available
  - current syntax does not allow it

packing_information(j) {	Descriptor	Value for jth atlas		
pin_codec_id[j]	u(8)	1 for HEVC main10, 3 for VVC main10		
pin_regions_count_minus1[j]	ue(v)	2		
for(i = 0; i <= pin_regions_count_minus1; i++) {		Looping over regions		
pin_region_tile_id[j][i]	u(8)	0	0	0
pin_region_type_id_minus2[j][i]	u(2)	2 (for attribute)	1 (for geometry)	1 (for geometry)
pin_region_top_left_x[j][i]	u(16)	0	0	960
pin_region_top_left_y[j][i]	u(16)	0	4640	4640
pin_region_width_minus1[j][i]	u(16)	1919	959	959
pin_region_height_minus1[j][i]	u(16)	4639	1159	1159
pin_region_unpack_top_left_x[j][i]	u(16)	0	0	0
pin_region_unpack_top_left_y[j][i]	u(16)	0	0	1160
pin_region_map_index[j][i]	u(4)	0	0	0
pin_region_rotation_flag[j][i]	u(1)	0	0	0
if(pin_region_type_id_minus2[j][i] + 2 == V3C_AVD || pin_region_type_id_minus2[j][i] + 2 == V3C_GVD)				
pin_region_auxiliary_data_flag[j][i]	u(1)	0	0	0
if(pin_region_type_id_minus2[j][i] + 2 == V3C_AVD) {				
pin_region_attr_type_id[j][i]	u(4)	0 (for texture)	-	-
pin_region_attr_partitions_flag[j][i]	u(1)	0	-	-
if(pin_region_attr_partitions_flag[j][i])				
pin_region_attr_partition_index[j][i]	u(5)	-	-	-
if(pin_region_attr_partition_index[j][i] == 0) pin_region_attr_partitions_minus1[j][i]	u(6)	-	-	-
}				
}		End looping over regions		
}				

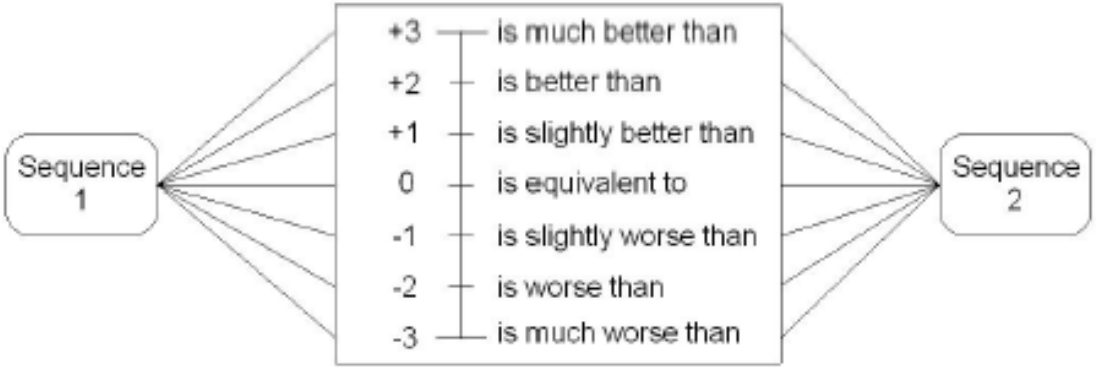
Packing information for V17 with 1 tile

Recommendations for MIV expert viewing sessions (m56321)

- Guidelines for expert viewing session
- Watching anchor/proposed video (10s) two times is enough
  Typical option: A/B/A/B, P.800 for instance
  Lowering frame rate is not acceptable
- Because matching the bitrate is hard, some flexibility is allowed
  Proponents can adjust texture and depth QP



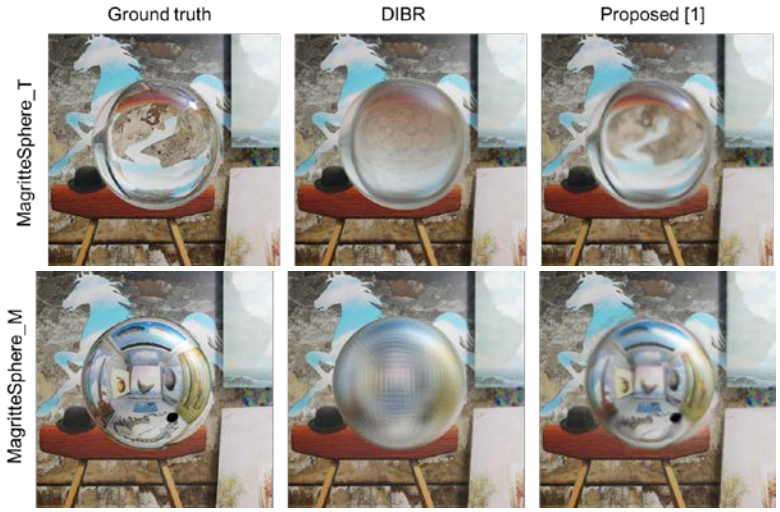
Example of viewing session



Sequence rating

Test Materials (m56450, 56632, 56730)

- Non-Lambertian CG content “MagritteSphere” by m56450
  - Rendered with Blender, two versions (transparent and mirror)
  - 1 frame, 21 × 21 views, 2000 × 2000 size
- NC content “Barn” by m56632
  - 97 frames, 5 × 3 perspective views, 1920 × 1080 size
  - Some holes exist (invalid depth)
  - Estimated depth by proprietary tool
- NC content “Breakfast” by m56730
  - Spec. is same as “Barn”
  - Captured in the dining room of Chateau de la Ballue



MagritteSphere sequence

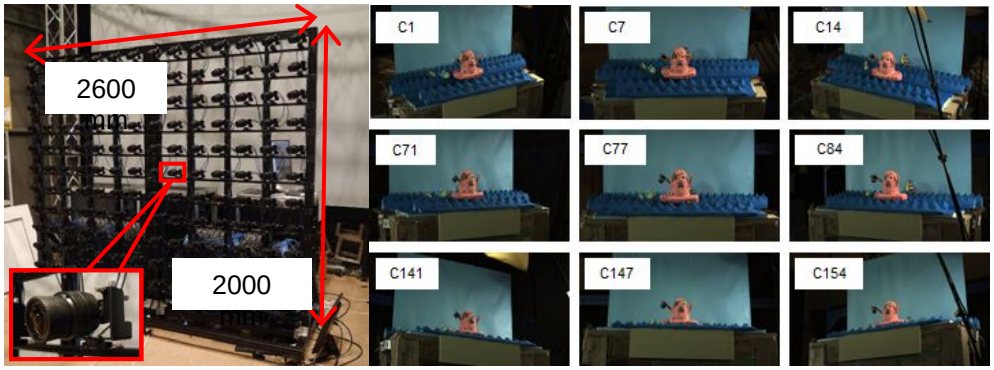
Barn sequence

Breakfast sequence

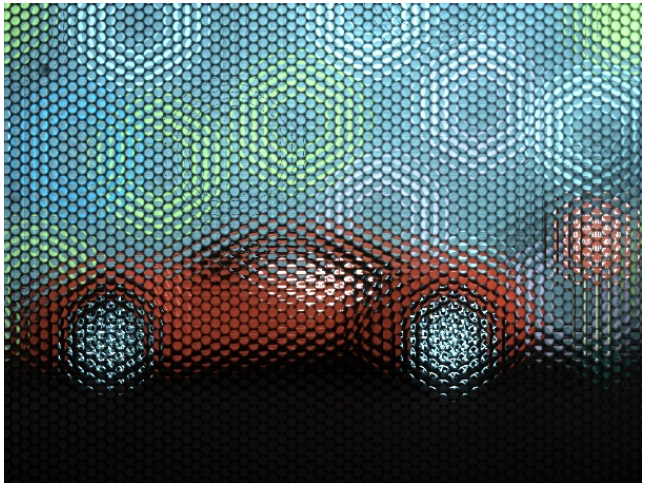
Test Materials (m56429, m56652, m56787)

- Natural content “RabbitStamp” by m56429
  - Captured by RayTrix R8 Plenoptic 2.0
  - 1 frame, 7 × 3 views, 2K × 2K size
- Natural content “Clay magic” by m56429
  - Captured by RayTrix R8 Plenoptic 2.0
  - 193-318 frames, 14 × 11 views, 1920 × 1080 size
  - Not parallel cameras -> good for testing MIV..?
- Natural content “Sports Car” by m56787
  - Captured by two plenoptic cameras
  - Recitification (rotate, crop, color matching)

RabbitStamp sequence



Clay magic sequence and capturing system



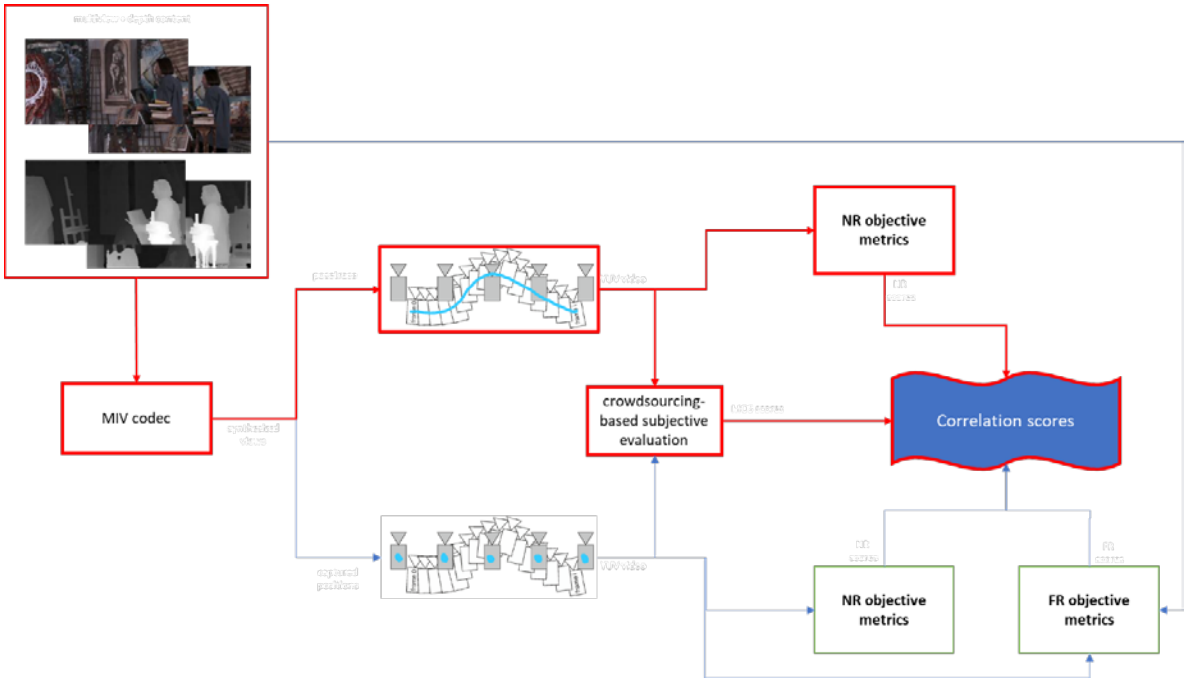
Left view of Sports Car sequence

Evaluation of objective quality metrics in MIV context (m56411)

- Current objective metrics are not suitable for MIV
  - Compression, synthesis artifacts
  - Evaluation on source views because GT is not enough (pose trace is needed)
  - Large differences between anchor and proposal
- Crowdsourcing-like subjective evaluation
- No-reference (NR) based on DL, but not re-trained for MIV
- Correlation with MOS: 0.76 RMSE, 0.75 PLCC, 0.73 SRCC

Category Rating		Predicted Quality/ (MOS)
Quality	Error Visibility	
Excellent	Imperceptible	5
Good	Perceptible Error, Not annoying	4
Fair	Visible error, Slightly annoying	3
Poor	Visible Error, Annoying	2
Bad	Visible error, Very annoying	1

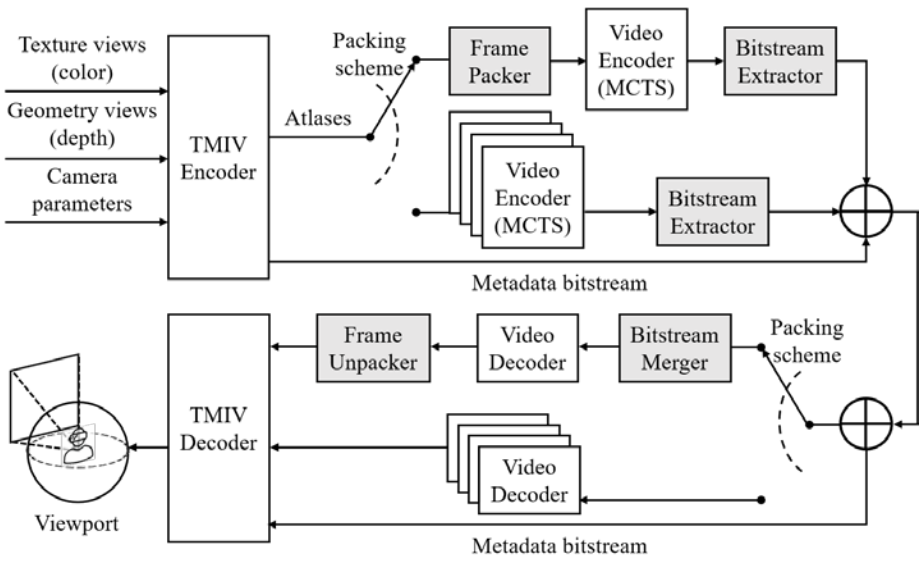
ACR scale score



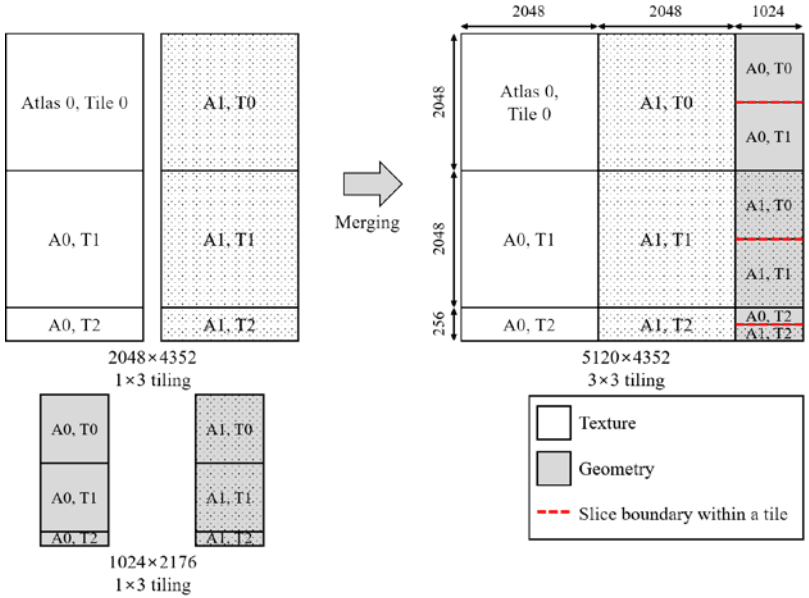
System overview of quality evaluation

Extraction and Merging on Frame Packed Video (m56591)

- SKKU MCSL proposed bitstream-level packing
- Fine-grained system in terms of the number of decoder instances
  - Current TMIV supports pixel-level frame packing -> provides high level two bitstreams
  - Proposed method can provide low level four or high level two bitstreams
  - 0.7%, 3.93% BD-rate saving because of deblocking filter deactivation
- Developed tile bitstream extracting/merging software on HEVC test model
- AhG recommendations
  - Tiling is needed for TMIV
  - Implementation of HLS to enable tiling on TMIV is more than welcome



System overview of packing method by MCSL



Example of bitstream-level packing

IEEE 3079

Thank You !

http://mcsl.skku.edu/
Questions > esryu@skku.edu