

JVET-AH0127 AHG9/AHG16: The SEI message design for scalable representation and layered reconstruction for generative face video compression

J. Chen, B. Chen, Y. Ye, R.-L. Liao (Alibaba)

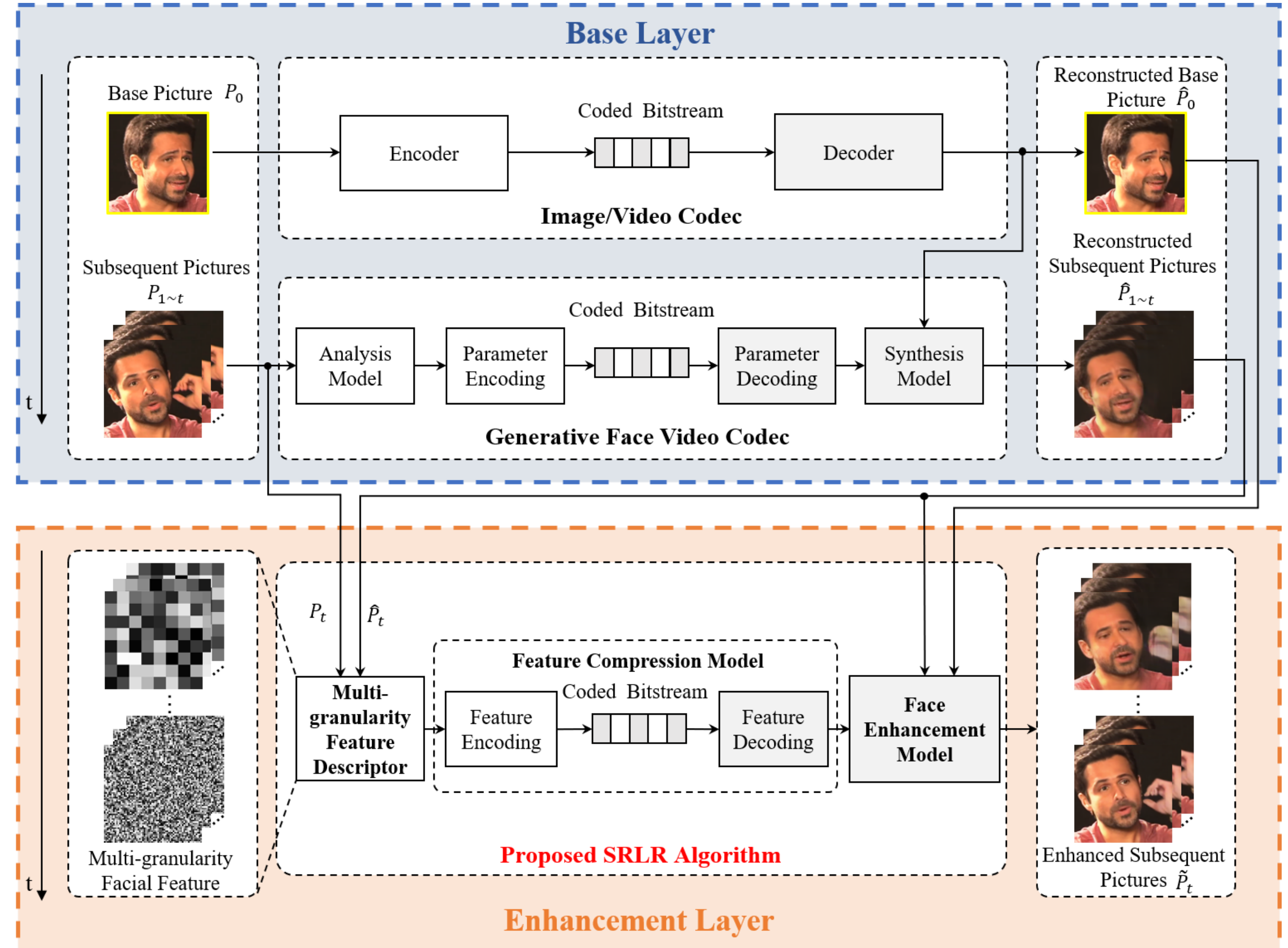
S. Yin, S. Wang (CityU)

Introduction

- Versions of the generative face video (GFV) SEI message were proposed in the previous meetings and was adopted to VSEI TuC draft 3 in the last meeting.
- In the current GFV SEI message, facial parameters of a variety of feature representations can be carried, which are used to generate the face pictures in the decoder side.
- To extend the bitrate and quality range, a scalable representation and layered reconstruction (SRLR) architecture was proposed for GFVC in JVET-AH0110 in this meeting.
- Thus, to support SRLR architecture of GFVC, the SEI message design for GFVC was extended in this contribution.

Overview of the SRLR architecture *(See more details in JVET-AH0110)*

- The current GFVC method is treated as the base layer
 - any face representation methods may be used
 - the facial parameters may be signaled with the current GFV SEI message
- An additional layer is added as the enhancement layer
 - multi-granularity feature (in format of matrix) representation



Proposal

- To support the enhancement layer features signaling, two options are provided in this contribution:
 - Option 1: extend the current GFV SEI message
 - Option 2: introduce a new SEI message named generative face video enhancement (GFVE) SEI message

Option 1: Extend the current GFV SEI message

- Additionally signal the enhancement layer face parameters in the GFV SEI message
 - Base layer parameters and enhancement layer parameters are signaled in the same SEI message
- A gating flag is introduced to control the signaling of the enhancement layer parameters
 - the enhancement layer parameters are optionally signaled
- An enhancement model “EnhancerNN()” is defined
 - the interface
 - the processing order
 - only invoked when enhancement layer parameters are present

Syntax changes of option 1

generative_face_video (payloadSize) {	Descriptor
gfv_id	ue(v)
gfv_cnt	ue(v)
...	
gfv_enhance_info_present_flag	u(1)
if(gfv_enhance_info_present) {	
gfv_enhance_matrix_element_precision_factor	ue(v)
gfv_num_enhance_matrice_minus1	ue(v)
for(i=0; i <= gfv_num_enhance_matrice_minus1; i++){	
gfv_enhance_matrix_height_minus1[i]	ue(v)
gfv_enhance_matrix_width_minus1[i]	ue(v)
for(j = 0; j <= gfv_enhance_matrix_height_minus1[i]; j++)	
for(k = 0; k <= gfv_enhance_matrix_width_minus1[i]; k++) {	
gfv_enhance_matrix_element[i][j][k]	ue(v)
if(!gfv_enhance_matrix_element[i][j][k])	
gfv_enhance_matrix_element_sign_flag[i][j][k]	u(1)
}	
}	
}	
}	
if(gfv_nn_present_flag)	
if(gfv_nn_mode_idc == 0) {	
while(!byte_aligned())	
gfv_nn_reserved_zero_bit_b	u(1)
for(i = 0; more_data_in_payload(); i++)	
gfv_nn_payload_byte[i]	b(8)
}	
}	
}	

Option 2: Introduce GFVE SEI message

- Define generative face video enhancement (GFVE) SEI message to signal the enhancement layer face parameters.
 - The enhancement model may be signaled or indicated
- Each GFVE SEI message is associated with a GFV SEI message
 - gfv_id and gfv_cnt of the associated GFV SEI message is signaled in the GFVE SEI message
- An enhancement model “EnhancerNN()” is defined
 - the interface
 - the processing order

Syntax of the proposed GFVE SEI message

generative_face_video_enhancement (payloadSize) {	Descriptor
gfve_id	ue(v)
gfve_gfv_id	ue(v)
gfve_gfv_cnt	ue(v)
gfve_nn_present_flag	u(1)
if(gfve_nn_present_flag) {	
gfve_nn_base_flag	u(1)
gfve_nn_mode_idc	ue(v)
if(gfve_nn_mode_idc == 1) {	
while(!byte_aligned())	
gfve_nn_reserved_zero_bit_a	u(1)
gfve_nn_tag_uri	st(v)
gfve_nn_uri	st(v)
}	
}	
gfve_matrix_element_precision_factor	ue(v)
gfve_num_matrices_minus1	ue(v)
for(i=0; i <= gfve_num_matrices_minus1; i++){	
gfve_matrix_height_minus1[i]	ue(v)
gfve_matrix_width_minus1[i]	ue(v)
for(j = 0; j <= gfve_matrix_height_minus1[i]; j++)	
for(k = 0; k <= gfve_matrix_width_minus1[i]; k++) {	
gfve_matrix_element[i][j][k]	ue(1)
if(!gfve_matrix_element[i][j][k])	
gfve_matrix_element_sign_flag[i][j][k]	u(1)
}	
}	
if(gfve_nn_present_flag)	
if(gfve_nn_mode_idc == 0) {	
while(!byte_aligned())	
gfve_nn_reserved_zero_bit_b	u(1)
for(i = 0; more_data_in_payload(); i++)	
gfve_nn_payload_byte[i]	b(8)
}	
}	
}	

Summary

- The SRLR architecture proposed in the JVET-AH0110 achieves a wider coverage of bitrate range for GFV.
- The design of the SEI message supporting SRLR architecture was proposed in this contribution.
 - Option 1: directly extend the current GFV SEI message
 - Option 2: introduce GFVE SEI message
- The GFV SEI message in TuC was implemented and tested in JVET-AH0239, which shows the feasibility of the SEI message solution for GFVC.
- Thus, it is suggested to include the proposed SEI message design to next draft of VSEI TuC to further support SRLR for GFVC.

Thanks