CROSS-PLANE CHROMA ENHANCEMENT FOR SHVC INTER-LAYER PREDICTION

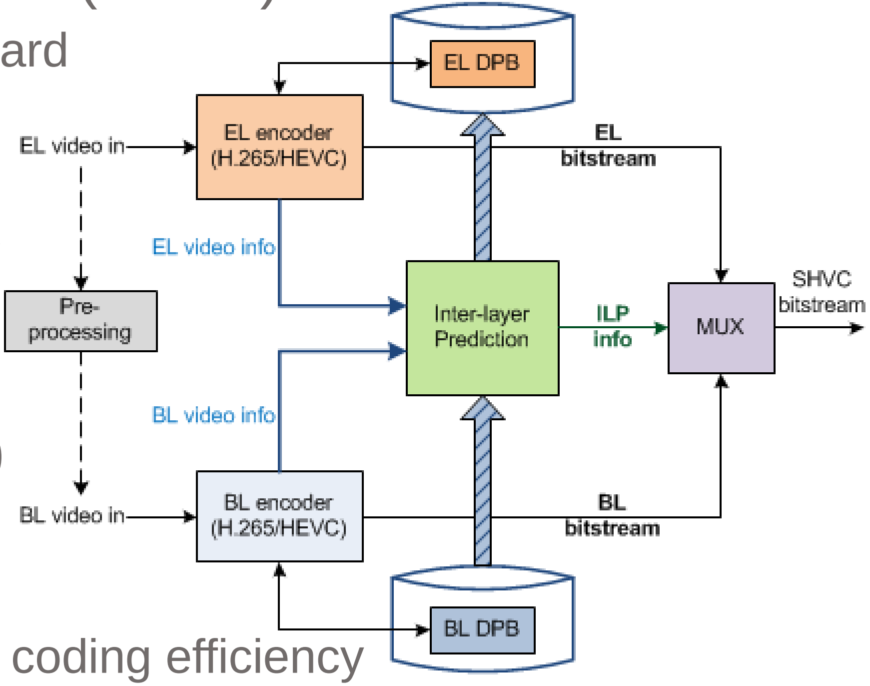
Jie Dong, Yan Ye, Yuwen He
InterDigital Communications, Inc.

PCS2013, Dec. 8-11, 2013, San Jose, USA

invention | collaboration | contribution

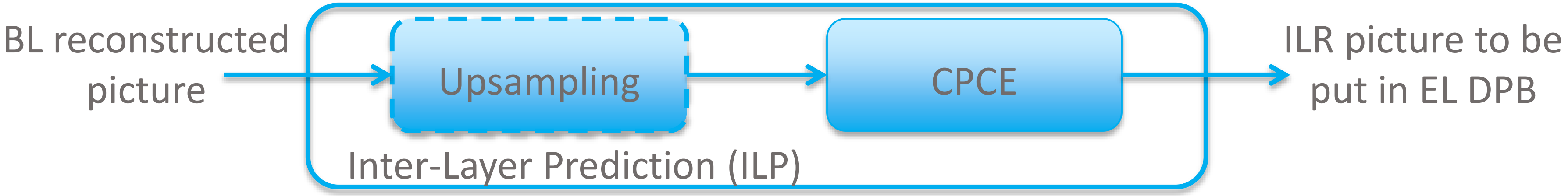
Introduction

- H.265/HEVC
  - New generation video coding standard finalized and approved in Jan. 2013.
  - Double the coding efficiency of its predecessor H.264/AVC, due to a number of new technical designs.
- Scalable extension of H.265/HEVC (SHVC)
  - On-going scalable video coding standard built on the top of H.265/HEVC.
  - Adopt multi-loop prediction structure: *processed BL reconstructed picture is inserted to EL DPB as an additional reference picture for EL coding*
- SHVC inter-layer prediction (ILP)
  - Process BL reconstructed picture (e.g., upsampling, enhancing, etc.)
  - Quality has a significant impact on EL coding efficiency



Cross-Plane Chroma Enhancement (CPCE)

- Performed in the ILP module, in order to improve the quality an inter-layer reference (ILR) picture's chroma planes

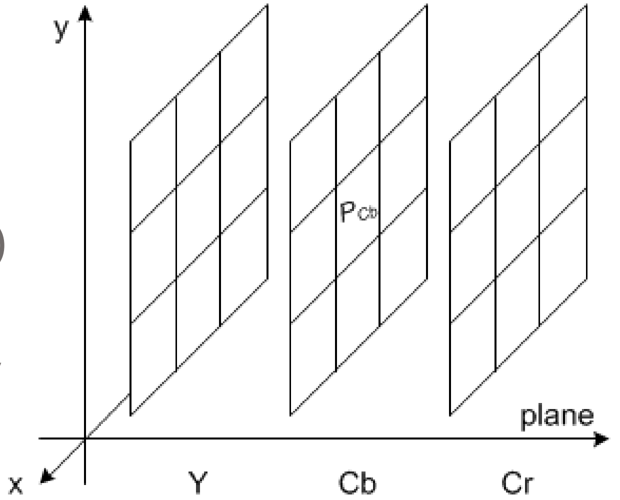


- General idea
  - Restore the blurred edges and textures in the chroma planes using the corresponding information from the luma plane
- Rationale
  - Cb and Cr planes have lower energy, and the edges and textures are too delicate to be preserved during quantization and compression.
  - Y plane, on the contrary, has higher energy and can maintain fair to good quality while QP is medium to high.
  - All three planes have high correlation in the edge structure.

Framework of Cross-Plane Chroma Enhancement (1)

- Cb pixel to be enhanced, P_{Cb}, is filtered by an optimal filter derived by LMMSE estimator.
- The optimal filter, $h_{Cb,opt}$, (usually 2-D) is extended to 3-D (2-D in spatial domain, 1-D in plane domain) because 2-D filter derived by LMMSE estimator has low-pass characteristics and cannot effectively restore the blurred edges in Cb plane.

$$Cb_{enh} = Y \otimes h_{Cb}(Y) + Cb \otimes h_{Cb}(Cb) + Cr \otimes h_{Cb}(Cr)$$



3-D support region to filter P_{Cb}

- Problem formulation

$$Diff\Big|_{h_{Cb}} = Y \otimes h_{Cb}(Y) + Cb \otimes h_{Cb}(Cb) + Cr \otimes h_{Cb}(Cr) - S_{Cb}$$

$$h_{Cb,opt} = \arg\min_{h_{Cb}} \sum_{y=0}^{H-1} \sum_{x=0}^{W-1} \left(Diff\Big|_{h_{Cb}}(x,y) \right)^2$$

- Not practical without simplification
  - High complexity (memory access and arithmetic operations)
  - Overhead for filter coefficient transition is unaffordable.

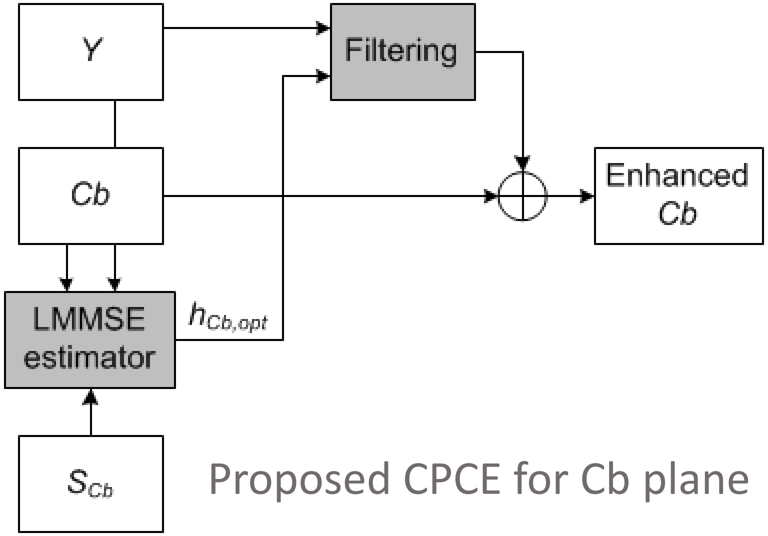
Framework of Cross-Plane Chroma Enhancement (2)

- Observations leading to simplification
  - $h_{Cb}(Cr)$ has little contribution to enhance a Cb pixel.
  - $h_{Cb}(Cb)$ has low-pass characteristics, and brings no performance penalty when reduced to an identity filter.
  - $h_{Cb}(Y)$ has high-pass characteristics, and cannot further improve the performance when the size is larger than 3×3.
- Proposed implementation

$$Cb_{enh} = Y \otimes h_{Cb}(Y) + Cb$$

$$Diff\Big|_{h_{Cb}} = Cb_{enh} - S_{Cb}$$

$$h_{Cb,opt} = \arg\min_{h_{Cb}} \sum_{y=0}^{H-1} \sum_{x=0}^{W-1} \left(Diff\Big|_{h_{Cb}}(x,y) \right)^2$$



Proposed CPCE for Cb plane

Only one 2-D high-pass filter applied on Y plane needs to be designed!

High-Pass Filter Design

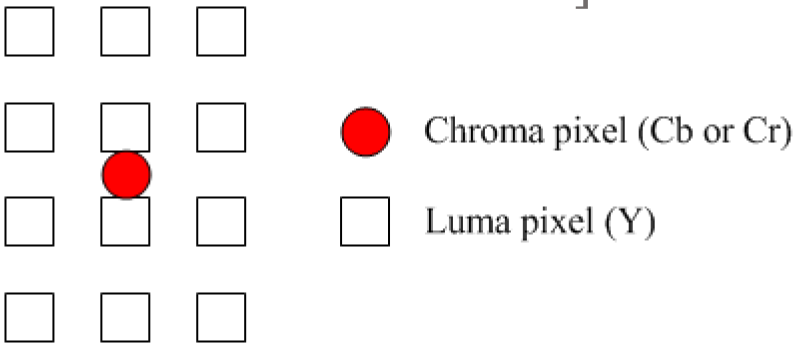
- Optimal high-pass filters designed for Cb plane (applied on Y plane), denoted as $h_{Cb,opt}$
  - Use LMMSE estimator to minimize the MSE between the original EL Cb plane and the enhanced Cb plane in the ILR picture

$$h_{Cb,opt} = \arg\min_{h_{Cb}} E\left[\left(\sum_{j=-2}^{1}\sum_{i=-1}^{1} h_{Cb}(i,j)Y(2x-i,2y-j) + Cb(x,y) - S_{Cb}(x,y)\right)^2\right]$$

- Optimal high-pass filters designed for Cr plane $h_{Cr,opt}$

$$h_{Cr,opt} = \arg\min_{h_{Cr}} E\left[\left(\sum_{j=-2}^{1}\sum_{i=-1}^{1} h_{Cr}(i,j)Y(2x-i,2y-j) + Cr(x,y) - S_{Cr}(x,y)\right)^2\right]$$

- With YUV 4:2:0 input, one chroma sample corresponds to a neighborhood of 3×4 luma samples, so the size of h_{Cb} and h_{Cr} is 3×4.



Quantization and Signaling

- $h_{Cb,opt}$ and $h_{Cr,opt}$ have real-valued coefficients, and therefore need to be quantized before transmission.
- Quantization:
  - 16-level uniform quantizer (Quantization stepsize denoted as Q_{step})

$$h_{Cb,opt}(i,j)/Q_{step} = f_{Cb}(i,j)$$

  $f_{Cb}(i,j)$ denotes the quantized coefficient and is an integer with the dynamic range from -8 to 7 (4-bit representation).

  - Q_{step} adapts to picture level, represented by two parameters Q_{Cb} and N_{Cb}

$$Q_{step} = Q_{Cb}/2^{N_{Cb}}$$

- Signaling:
  - Slice header (61 bits for each chroma plane)
    - 1-bit flag: indicating On/Off for certain chroma plane
    - 11 filter coefficients explicitly coded: 4 bits each (the remaining one is implicitly derived using zero-summation constraint)
    - Q_{Cb} : 11 bits (10 bits for magnitude and 1 bit for sign)
    - N_{Cb} : 5 bits

Process of CPCE

- To enhance a Cb pixel located at position (*x*,*y*), denoted as *Cb*(*x*, *y*)

1. High-pass filter f_{Cb} is first applied to the corresponding 3×4 luma neighboring pixels to generate the intermediate result *z*(*x*, *y*)

$$z(x,y)=\sum_{j=-2}^{1}\sum_{i=-1}^{1} f_{Cb}(i,j)Y(2x-i,2y-j)$$

2. *z*(*x*, *y*) is scaled by $Q_{Cb}/2^{N_{Cb}}$ to the normal range, denoted as *o*(*x*, *y*). The integer realization of this scaling step is shown below.

$$o(x,y)=\mathrm{Sign}(Q_{Cb}z(x,y))\left((\mathrm{Abs}(Q_{Cb}z(x,y))+2^{N_{Cb}-1})\gg N_{Cb}\right)$$

3. The offset value *o*(*x*, *y*) is then added to *Cb*(*x*, *y*) to obtained the enhanced Cb pixel, denoted as $Cb_{enh}(x,y)$

$$Cb_{enh}(x,y)=Cb(x,y)+o(x,y)$$

Test Conditions

- SHVC common test conditions [1] specified by JCT-VC is used.

Software platform	SHM-1.0
Test sequences	2 Class A: 2560×1600 5 Class B: 1080*p*
Type of scalabilities	Spatial scalability (1.5x, 2x) SNR scalability
Configurations	All Intra (AI) Random Access (RA) Low Delay P (LDP) Low Delay B (LDB)
QP for spatial scalability	BL *QP* (22, 26, 30, 34) EL QP = BL $QP + \Delta QP$ (ΔQP = 0, 2)
QP for SNR scalability	BL *QP* (26, 30, 34, 38) EL QP = BL $QP + \Delta QP$ (ΔQP = -6, -4)
Performance measurement	BD-rate [2]

[1] X. Li, J. Boyce, P. Onno, Y. Ye, "Common SHM test conditions and software reference configurations," JCT-VC document, JCTVC-L1009, Geneva, Switzerland, Jan. 2013.

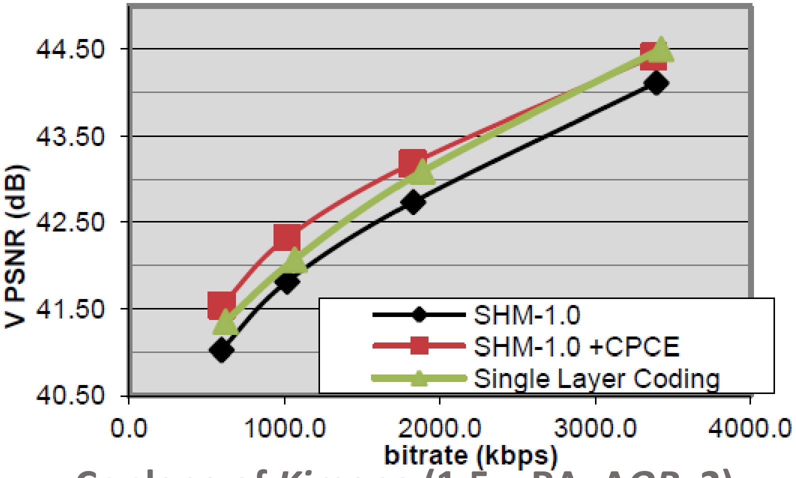
[2] G. Bjontegaard, "Calculation of average PSNR differences between RD-curves," document VCEG-M33, ITU-T SG16/Q6, Apr. 2001.

R-D Performance

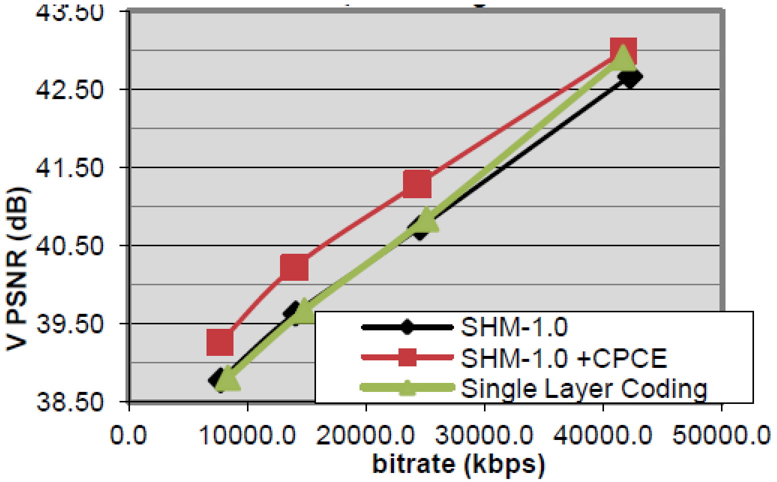
BD-rate averaged over sequences and *QPs*

	AI 2x			AI 1.5x					
	Y	Cb	Cr	Y	Cb	Cr			
Class A	-0.9	-7.8	-6.2						
Class B	-0.8	-6.4	-8.5	-0.7	-8.2	-10.4			
Average	-0.8	-6.8	-7.8	-0.7	-8.2	-10.4			
	RA 2x			RA 1.5x			RA SNR		
	Y	Cb	Cr	Y	Cb	Cr	Y	Cb	Cr
Class A	-0.5	-11.2	-7.7				-0.4	-9.0	-5.4
Class B	-0.3	-8.0	-9.7	-0.3	-10.5	-12.1	-0.3	-7.7	-8.6
Average	-0.4	-9.0	-9.1	-0.3	-10.5	-12.1	-0.3	-8.0	-7.7
	LDP 2x			LDP 1.5x			LDP SNR		
	Y	Cb	Cr	Y	Cb	Cr	Y	Cb	Cr
Class A	-0.2	-9.3	-5.2				-0.3	-7.3	-4.4
Class B	-0.2	-4.6	-5.8	-0.2	-7.6	-10.1	-0.2	-5.2	-6.1
Average	-0.2	-5.9	-5.6	-0.2	-7.6	-10.1	-0.2	-5.8	-5.6
	LDB 2x			LDB 1.5x			LDB SNR		
	Y	Cb	Cr	Y	Cb	Cr	Y	Cb	Cr
Class A	-0.2	-9.7	-5.8				-0.3	-8.0	-5.0
Class B	-0.2	-5.1	-6.7	-0.2	-8.2	-10.8	-0.2	-5.6	-6.9
Average	-0.2	-6.4	-6.4	-0.2	-8.2	-10.8	-0.2	-6.3	-6.3

Examples of R-D curves



Cr plane of *Kimono* (1.5x, RA, Δ*QP*=2)



Cr plane of *ParkScene* (1.5x, AI, Δ*QP*=2)

Performance Comparison vs. Simulcast

- Currently in SHM-1.0
  - Chroma performance lags luma performance 6.5% to 10.2% for RA/LDP/LDB
- SHM1.0 + CPCE
  - Comparable performance for luma and chroma components

	SHM-1.0			SHM-1.0 + CPCE		
	Y	Cb	Cr	Y	Cb	Cr
AI	-26.8%	-26.4%	-26.7%	-27.4%	-31.6%	-32.7%
RA	-20.6%	-12.2%	-10.4%	-20.9%	-20.1%	-18.5%
LDP	-15.4%	-8.1%	-6.2%	-15.6%	-13.9%	-12.5%
LDB	-14.7%	-8.2%	-6.3%	-14.9%	-14.4%	-13.2%

Conclusion

- Currently, the ILP of chroma in SHVC is relatively weak.
- Propose the CPCE scheme to enhance the chroma planes of ILR pictures for SHVC.
  - Blurred edges and textures in the chroma planes are restored by using the corresponding edge information from the luma plane.
  - Related problems, such as quantization, signaling, and complexity reduction are well addressed.
- Performance
  - Compared with SHM-1.0, BD-rate reductions are -7.5% and -8.5% for the Cb and Cr planes, respectively (averaged over all sequences and test conditions)
  - The coding efficiency of chroma components catches up with that of the luma component.