Improving SHVC Performance with a Joint Layer Coding Mode

Xiem HoangVan, João Ascenso, and **Fernando Pereira**

Instituto Superior Técnico – Instituto de Telecomunicações, Lisbon, Portugal

hoang.xiem@lx.it.pt, joao.ascenso@lx.it.pt, fp@lx.it.pt

Shanghai, March 23rd , 2016

Outline

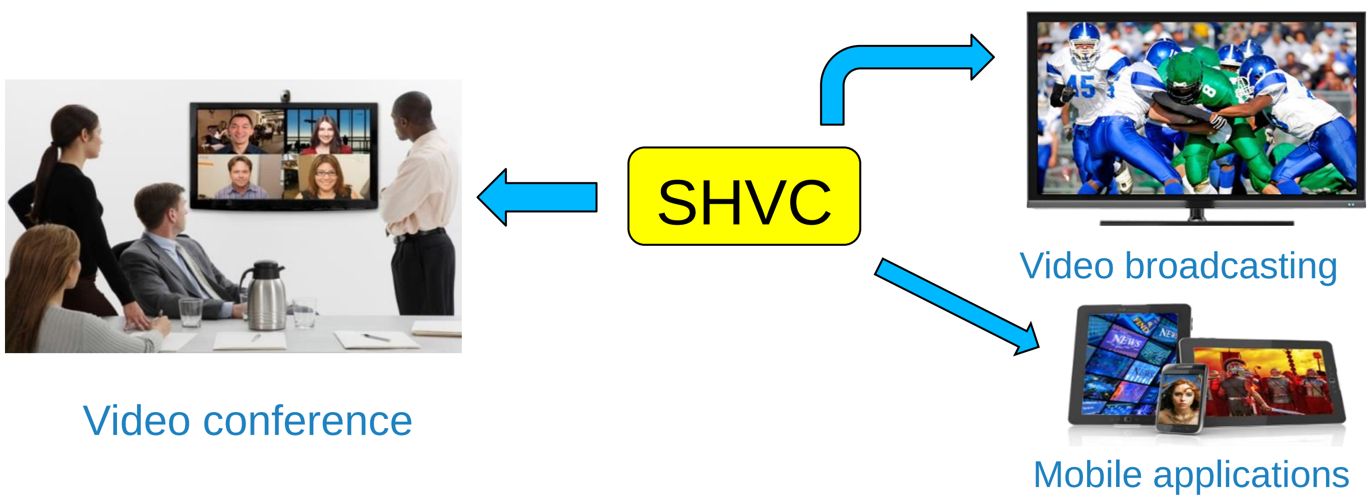
1. Motivation and Objective
2. Background Work
3. Proposed Joint Layer Coding Mode
4. Performance Evaluation
5. Conclusions

Motivation

- ❖ The heterogeneity of networks and communication devices asks for an efficient scalable video coding engine

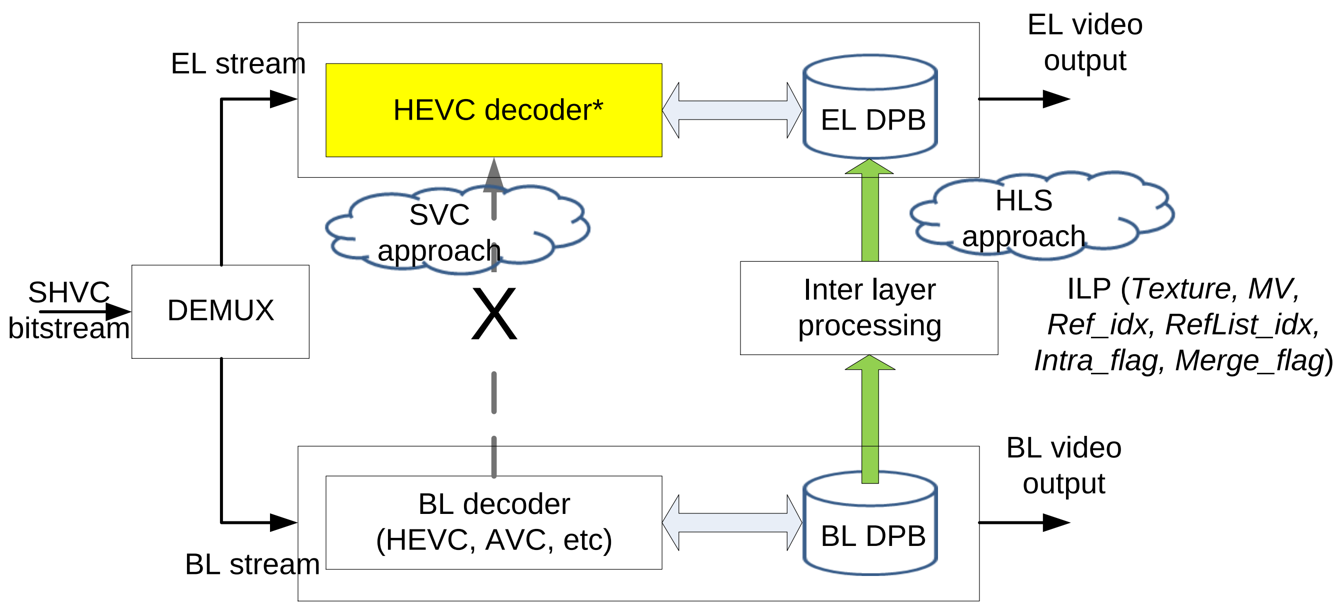
- ❖ The HEVC increased compression efficiency asks for a novel scalable video coding solution

→ HEVC Scalable extension (**SHVC standard**) may be a solution!



Objective

Improving SHVC compression efficiency by proposing a novel Joint Layer Coding Mode

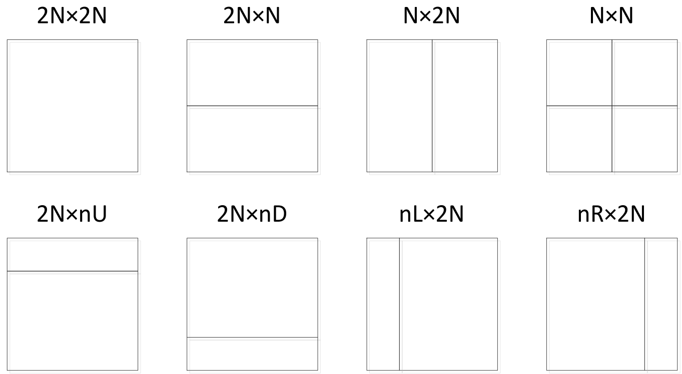


SHVC decoder architecture

Outline

SHVC Prediction Modes

- Prediction Unit (PU) in SHVC is defined to efficiently code a coding unit (CU) using either **Inter prediction** or **Intra prediction** coding modes
- A CU can be split into 8 partitions



- SHVC Intra prediction considers 35 prediction modes
- SHVC Inter prediction includes the traditional motion estimation (ME) and the new Merge mode
- SHVC Merge mode candidates includes the **spatial**, **temporal** and **inter layer candidates**

Related Works

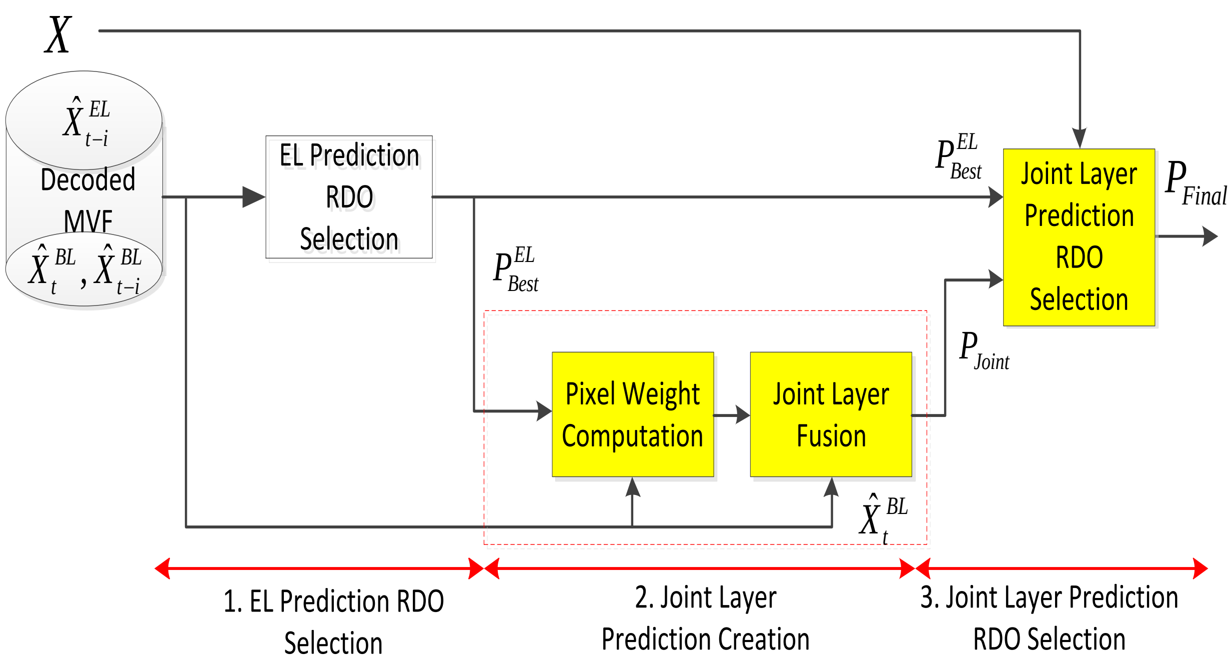
Most recent research focus on *inter-layer processing* to create additional EL references to improve the SHVC compression efficiency such as:

Year	Authors	Proposals	Venues
2013	Xiang *et al.*	Generalized inter-layer residual prediction	ICIP
	Lai *et al.*	Combined temporal and inter-layer prediction	PCS
	Guo *et al.*	Wiener filter is adaptively applied to BL decoded frames	
	Lai *et al.*	Directional filter is adaptively applied to BL decoded frames	
2014	Laude *et al.*	Scalable extension of HEVC using enhanced inter-layer prediction	ICIP
	Aminlou *et al.*	Differential coding using enhanced inter-layer reference picture for scalable extension of H.265/HEVC video codec	TCSVT
2015	Xiem *et al.*	Improving enhancement layer merge mode for HEVC scalable extension	PCS

Outline

Proposed EL Coding Mode Selection Architecture

The basic target of the proposed joint layer prediction creation is to obtain better CU predictions than with traditional EL coding modes



Joint Layer Prediction Creation

The proposed joint layer prediction creation includes two main steps:

(1) Joint layer fusion:

$$P_{Joint}\big(x, y, w(x, y)\big) = w(x, y) \times P_{Best}^{EL}(x, y) + \big(1 - w(x, y)\big) \times \hat{X}_t^{BL}(x, y)$$

Here,

$P_{Joint}\big(x, y, w(x, y)\big)$: Joint layer prediction

$w(x, y)$: Weighting term

$P_{Best}^{EL}(x, y)$: EL traditional best prediction

$\hat{X}_t^{BL}(x, y)$: BL reconstruction

(2) Pixel weight computation

Pixel Weight Computation

- **Ideal solution**: Using the square difference, SD_{BL}, between the **original information**, X_t, and $\hat{X}_t^{BL}$ and the square difference, SD_{EL}, between the X_t and P_{Best}^{EL}

However, the overhead bits associated to the weight of each pixel is too heavy !

- **Proposed weight computation**: Exploiting only the available decoded information

→ ***No bitrate overhead is required***

→ ***the weight computation can be synchronously performed at both encoder and decoder***

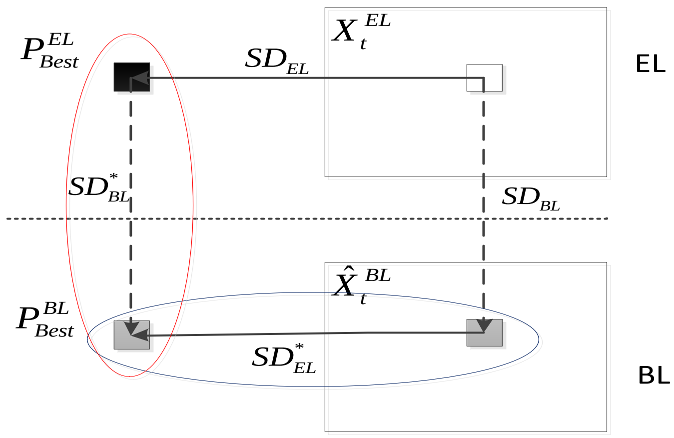
Proposed Weight Computation

❖ The proposed weight computation proceeds with the following steps:

- **Step 1: Pixel weight initialization**: First, the square differences for each pixel, SD^*_{BL} and SD^*_{EL} are computed as:

$$SD^*_{BL}(x,y) = \left(P^{EL}_{Best}(x,y) - P^{BL}_{Best}(x,y)\right)^2$$

$$SD^*_{EL}(x,y) = \left(\hat{X}^{BL}_t(x,y) - P^{BL}_{Best}(x,y)\right)^2$$



Then**, initial weight is computed** as:

$$w_{ini}(x,y) = \frac{SD^*_{BL}(x,y) + 1}{SD^*_{BL}(x,y) + SD^*_{EL}(x,y) + 2}$$

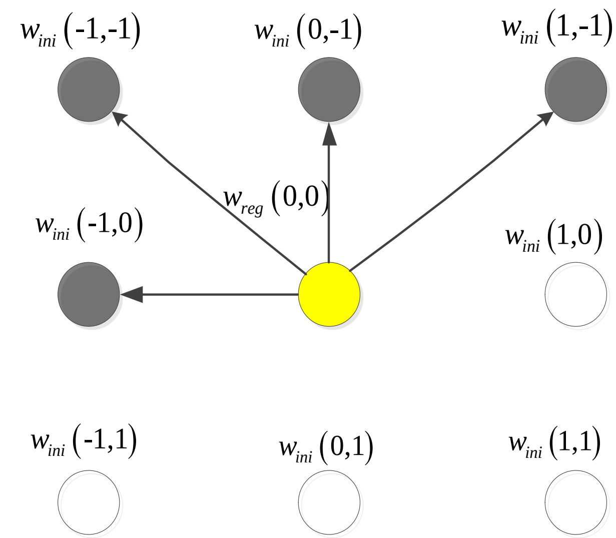
- **Step 2: Pixel weight regularization:** To further improve the weight accuracy

Pixel Weight Regularization

Key idea: Using the spatial neighborhood pixels to regularize the initial weight

(1) Weight candidates definition: A weight candidate list, W_List, for each pixel is defined as:

$$W\_List = \{w_{ini}(i,j)\}\ ;\ (i,j) = \{-1,0,1\}$$



(2) Joint layer prediction candidates:

The joint layer prediction associated to each weight candidate is computed as:

$$P(x,y,w_{ini}(i,j)) = w_{ini}(i,j) \times P_{Best}^{EL}(x,y) + (1 - w_{ini}(i,j)) \times \hat{X}_t^{BL}(x,y)$$

Pixel Weight Regularization (cont.)

(3) Spatial coherence measurement definition: a spatial coherence metric is defined as the sum of square differences between the joint layer prediction, $P\big(x, y, w_{ini}(i,j)\big)$, and its four "reliable" neighboring pixels, $P\big(m, n, w(m,n)\big)$ with $(m,n) \in P\_List$ and P_List is:

$$P\_List = \{(-1,0); (-1,-1); (0,-1); (1,-1)\}$$

The spatial coherence metric is computed as:

$$SSD\big(x, y, w_{ini}(i,j)\big) = \sum_{(m,n)\in P\_List} \Big(P\big(x, y, w_{ini}(i,j)\big) - P\big(m, n, w(m,n)\big)\Big)^2$$

Pixel Weight Regularization (cont.)

(4) Regularized weight creation: The regularized weight for each pixel is then obtained by selecting the weight candidate that minimizes $SSD\left(x, y, w_{ini}(i, j)\right)$ as follows:

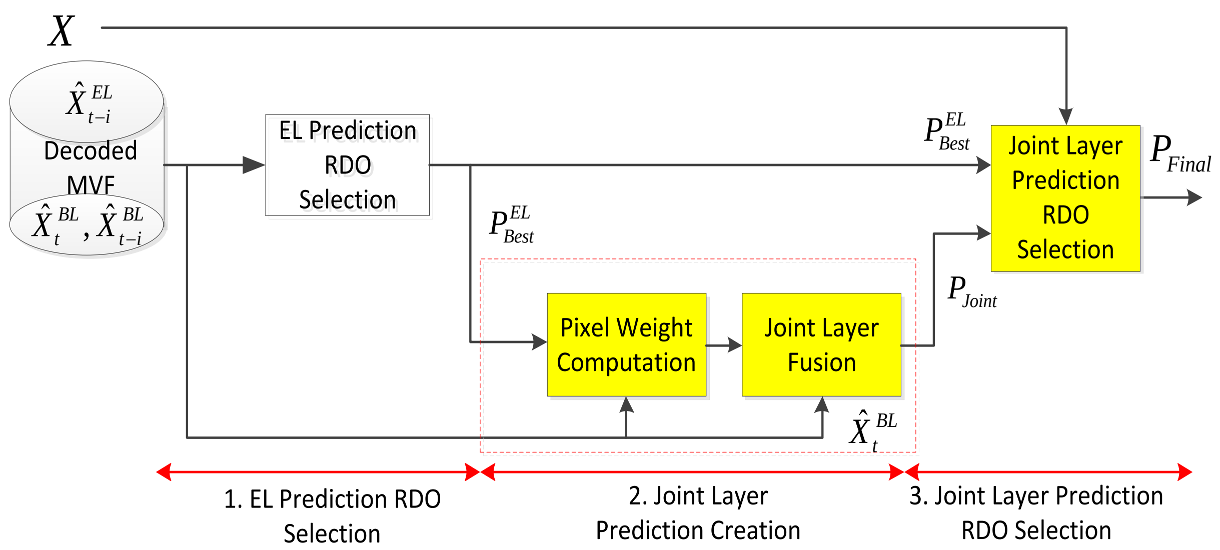
$$w_{reg}(x, y) = \underset{w_{ini}(i) \in W\_List}{\operatorname{argmin}} SSD\left(x, y, w_{ini}(i, j)\right)$$

Finally, the regularized weights are used to create the joint layer prediction for each pixel as:

$$P_{Joint}\left(x, y, w(x, y)\right) = w_{reg}(x, y) \times P_{Best}^{EL}(x, y) + \left(1 - w_{reg}(x, y)\right) \times \hat{X}_t^{BL}(x, y)$$

Joint Layer Prediction RDO Selection

- Perform a selection between the proposed joint layer prediction and the traditional EL best prediction under a RDO mechanism
- Require a flag indicating the final selected mode



Outline

Test Conditions and Benchmarks

❖ Test conditions

Sequences	Spatial resolution	Temporal resolution	Number of test frames
RaceHorses	416×240	30 Hz	297
BlowingBubbles		50 Hz	497
BasketballPass		50 Hz	497
PartyScene	832×480	50 Hz	497
BQMall		60 Hz	600
GOP size	2, 4 (LD-P, LD-B), 8 (RA)		
Quantization parameters	$QP_{BL} = 34$ $QP_{EL} = \{32; 30; 28; 26\}$		

❖ Benchmarks:

- SHVC standard with the conventional coding modes
- SHVC with improved EL merge mode [11]

BD-Rate Savings with the proposed Joint Layer Mode

Sequences	GOP 2		GOP 4 (LD-P)	
	Ref [11]	Proposed	Ref [11]	Proposed
RaceHorses	-3.27	-4.04	-0.32	-3.44
BlowingBubbles	-3.49	-4.06	-2.14	-3.76
BasketballPass	-3.01	-3.86	-0.33	-3.22
PartyScene	-2.62	-3.12	-1.85	-3.26
BQMall	-2.93	-3.88	-2.12	-4.30
Average BD-Rate to SHVC	**-3.06**	**-3.79**	**-1.35**	**-3.60**
Average BD-Rate to [11]		**-0.73**		**-2.24**

Sequences	GOP 4 (LD-B)		GOP 8 (RA)	
	Ref [11]	Proposed	Ref [11]	Proposed
RaceHorses	-0.64	-2.69	-0.21	-1.79
BlowingBubbles	-1.90	-2.78	-2.36	-3.64
BasketballPass	-0.45	-2.06	-0.08	-1.85
PartyScene	N	N	-1.51	-2.57
BQMall	N	N	-1.19	-3.11
Average BD-Rate to SHVC	**-1.00**	**-2.51**	**-1.07**	**-2.59**
Average BD-Rate to [11]		**-1.51**		**-1.52**

Proposed SHVC extension vs. SHVC standard

- SHVC extension with the novel joint layer coding mode always outperforms SHVC with the standard prediction modes for all test sequences and all GOP sizes

- The higher gains are obtained for the smaller GOP sizes as the temporal distance between the current and the reference pictures is smaller; thus, the weight computed with the proposed solution is more accurate for case of small GOP size

Proposed SHVC extension vs. SHVC with improved Merge mode [11]

- SHVC extension with the novel joint layer coding mode always outperforms SHVC with the improved Merge mode for all GOP sizes
- The proposed SHVC extension brings a significant compression efficiency gains not only for low motion sequence but also for high motion sequences
- The proposed SHVC extension also requires lower processing complexity than the improved Merge mode solution in [11] due to the absence of the high complexity motion refinement process

Outline

5. Conclusions

- ❖ This paper proposes a novel **joint layer coding mode** for the SHVC standard jointly exploiting the EL and BL decoded information
- ❖ The proposed SHVC with joint layer coding mode outperforms the standard SHVC solution, notably by up to 4.3% BD-Rate saving
- ❖ Future work may consider to improve the accuracy of the fusion weight to better create the joint layer prediction quality