Low-Complexity Spatial Scalability Scheme Using HEVC for 4K and VR Videos

Glenn Herrou, W. Hamidouche, L. Morin

Low-Complexity Spatial Scalability Scheme using HEVC for 4K and VR Videos

Glenn Herrou*, Wassim Hamidouche†,* and Luce Morin†,*

*IRT b<>com, Cesson-Sevigne, France †IETR/INSA Rennes, France

glenn.herrou@b-com.com {whamidou,lmorin}@insa-rennes.fr

Scalable video coding enables to compress the video at different formats within a single layered bitstream. Scalable High efficiency Video Coding (SHVC), the scalable extension of the High Efficiency Video Coding (HEVC) standard, enables x2 spatial scalability, among other additional features. The closed-loop architecture of the SHVC codec is based on the use of multiple instances of the HEVC codec to encode the video layers, which considerably increases the encoding complexity. As part of the Joint Video Exploration Team (JVET) effort, E. Thomas et al. proposed a new scalable scheme [1], based on a polyphase sub-sampling performed prior to encoding, achieving x2 spatial scalability with a single HEVC encoder instance, thus greatly reducing the coding complexity compared to SHVC.

The polyphase sub-sampling technique takes one over four pixels, in each 2x2 block of the input image (luma and chroma planes), to create four resolution components. For the 4:2:0 format, the polyphase decomposition introduces a misalignment between the chroma pixels of the different sub-resolution images. Thus, when a frame from a different resolution component is used as reference for an inter-prediction, the derived chroma motion vector is inherently wrong. In this paper, we propose a filter to realign the chroma positions of the different sub-images, achieving an average 14% BD-rate gain for both chroma planes compared to the original polyphase decomposition.

We also propose to replace the polyphase sub-sampling by a wavelet-based decomposition in order to avoid the potential aliasing introduced in the four polyphase sub-sampled images. The proposed wavelet-based decomposition process is as follows: an integer-to-integer discrete wavelet transform (Haar or Le Gall 5/3) is applied on the full resolution input signal, the LL sub-band is used as base layer and the LH, HL and HH sub-bands form the enhancement layer.Then, to make the decomposed signal suitable for a standard HEVC encoder, we add the LL sub-band to each of the three high-frequency sub-bands, thus enabling inter-layer predictions.

The proposed decompositions have been evaluated on a set of 4K sequences using HEVC reference software (HM16.12) with a random access configuration (GOP16 and 1 sec. intra period). Compared to SHVC, average BD-rate gains of -3.2%, -6.4% and -2.3% for the luma plane and average BD-rate losses of 57%, 32% and 42% for both chroma planes are achieved for the polyphase with chroma alignment, Haar and Le Gall 5/3 decompositions, respectively. In addition, the proposed scalable coding chain shows a 50% complexity reduction at both encoding and decoding sides.

References

[1] E. Thomas, "Polyphase Subsampled Signal for Spatial Scalability," *Document JVET-B0043*, San Diego, February 2016.