Joint Layer Prediction for Improving SHVC Compression Performance and Error Concealment

Xiem HoangVan and Byeungwoo Jeon, *IEEE Senior Member*

***Abstract*—Scalable High Efficiency Video Coding (SHVC) standard is expected to play a more important role in the heterogeneous landscape of broadcasting, multimedia, networks, and various services applications as it is specified as a layered coding technique in the ATSC (Advanced Television Systems Committee) 3.0. However, its block-based structure of temporal and spatial prediction makes it sensitive to information loss and error propagation due to transmission errors. In this context, we propose an improved SHVC with a joint layer prediction (JLP) solution which adaptively combines the decoded information from the base and the enhancement layers to create an additional reference for the SHVC enhancement encoder. To optimize the quality of the joint prediction, the minimum mean square error (MMSE) estimation is executed in computing a combination factor which gives weights to each contribution of the decoded information from the layers. In addition, the proposed JLP is integrated into the SHVC decoder to work as an error concealment solution to mitigate the error propagation happening inevitably in practical video transmission. Experiments have shown that the proposed SHVC framework significantly outperforms its relevant benchmarks, notably by up to 14.8% in bitrate reduction with respect to the standard SHVC codec. The proposed SHVC error concealment strategy also greatly improves the concealed picture quality as well as reducing the problem of error propagation when compared to conventional error concealment approaches.**

***Index Terms*—Compression efficiency, error concealment, joint layer prediction, MMSE, scalable video coding**

I. INTRODUCTION

NOWADAYS, video contents can be stored or delivered in several formats in quality, frame rate, spatial resolution, chroma sub-sampling ratio, or bit-depth in order to cover a wide range of user's requirements associated most probably with available bandwidth and/or equipment resources like memory or computation as well as the quality requested by the user. Quite expectedly, encoding and delivering the video in such a way calls for scalable encoding scheme unless considerable increases of resources in both storage and bandwidth can be borne with. In addition, the increasing diversity of heterogeneous and dynamic transmission environments have expedited the development of a new scalable video coding standard to achieve a far more flexible, easy to implement solution over existing platforms but still having better compression performance than the prior standard of H.264/SVC [1]. It has led the experts group to make MPEG HEVC/ITU-T H.265 to further standardize the scalable extension of HEVC (known as SHVC) [2, 3].

Note that the ATSC 3.0 has considered several broadcasting scenarios, and identified the scalable video as an important functionality for combined services of HD/UHD and/or fixed/mobile services [4, 5, 6, 7]. Fig. 1 shows an example of such broadcasting scenario in which a SHVC scalable compressed bitstream consisting of a base layer (BL) and an enhancement layer (EL) provides a scalable broadcasting service at the same time over different environments. The SHVC scalable bitstream can be broadcasted over the MPEG-2 TS (transport stream) [8] or through the Multi-input Multi-output (MIMO) broadcasting scheme [9]. The user can receive and decode only the base layer or both layers depending on its display, energy, and network capabilities.

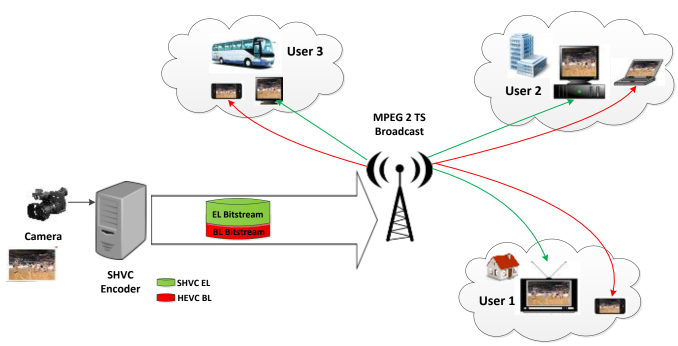

Fig. 1. An example of ATSC 3.0 broadcasting scenario

SHVC employed the traditional layered coding approach again but in a slight different way. That is, different from H.264/SVC [1], the coding tools of existing HEVC (which is used as the BL) are kept intact and the scalable functionality is implemented by defining additional signaling capabilities at macroblock-level to indicate whether EL macroblock is predicted from BL or other EL layers. In SHVC, the reconstructed BL picture, taken as an inter-layer reference (ILR) picture, is put to the EL prediction buffer following the normal inter-layer processing [3, 10]. The reference index signaling mechanism in (non-scalable version of) HEVC still works perfectly to identify whether the prediction for EL block

Manuscript received March 31st, 2018, revised July 10th and October 20th, 2018, accepted October 28th

X. HoangVan is with the VNU-University of Engineering and Technology, Hanoi, Vietnam (xiemhoang@vnu.edu.vn)

B. Jeon is with Sungkyukwan University, South Korea (bjeon@skku.edu)