An online SVM based side information creation for efficient distributed scalable video coding

Xiem HoangVan[1], Thao Nguyen Thi Huong[2],
[1] VNU-University of Engineering and Technology
[2] Posts and Telecommunications Institute of Technology

***Abstract* - With the significant increase of the network heterogeneity and the wide use of emerging video applications such as wireless sensor networks, video surveillance systems or remote sensing, the Distributed Scalable Video Coding (DSVC) is a potential solution for efficiently transmitting and storing video data due to its high compression efficiency and low encoding complexity capabilities. In DSVC framework, Side Information (SI), created at the decoder side by exploiting the temporal and inter-layer correlations between decoded frames, plays an important role as it directly affects to the final DSVC coding performance. Therefore, this paper proposes a novel SI creation solution which explicitly formulates the SI creation as a classification problem and employs an online learning Support Vector Machine (SVM) engine to fuse several SI candidates. Experiments conducted for a rich set of test sequences show that the proposed SI creation solution significantly outperforms the previous DSVC SI creation methods in terms of SI quality while slightly introducing the computational complexity.**

I. INTRODUCTION

The growing heterogeneity and dynamic nature of the networks, terminals, and usage environments in Internet of Things (IoTs) era has boosted the need for powerful video coding engines which are able to efficiently adapt to changing consumption conditions. The emerging video applications such as video surveillance systems, wireless visual networks or remote sensing [1] are typically asking for a scalable video coding solution with not only the compression efficiency but also the low complexity and error resiliency capabilities. To address these requirements, a novel Distributed Scalable Video Coding framework has been recently introduced [2].

Distributed Scalable Video Coding (DSVC) refers to all coding solutions which use distributed source coding principles [1] in coding the video data with different types of scalabilities, e.g., temporal, spatial and/or quality. In such framework, a layered coding approach was adopted with a base layer (BL) and one or several enhancement layers (ELs). While the HEVC standard [3] is used to compress the BL video data, the distributed coding solution is used to compress the ELs data to make use of its low complexity and error resilience benefits.

In DSVC framework, Side Information (SI) plays an important role due to its direct contribution to the quality of the final reconstruction frame. In DSVC, SI creation is usually performed at the decoder to reconstruct the EL frames after knowing the correlation between the encoder EL and the decoder SI [2].

Many SI creation solutions have been proposed in the literature so far, notably based on frame interpolation [4] and extrapolation [5]. Other solutions create the SI at the decoder in a tentative way [6] or rely on auxiliary data received from the encoder for the parts of the image that are more difficult to estimate [7]. Moreover, the initial SI estimation can be refined based on the data already decoded for each frame, e.g. the successive DCT coefficients [8]. Moreover, hybrid approaches combining these elementary approaches can be also used. While the above SI creation solutions can be directly employed in a DSVC framework but this is not an efficient solution because, in scalable video coding, not only the EL decoded forward and backward frames but also the decoded BL frame and its motion information can be exploited.

In the recent DSVC SI creation solutions [2, 9], several SI candidates were created before combining in a fusion mechanism. The SI quality is naturally driven by the efficiency of the fusion solution. Hence, the fusion stage is a critical problem. In this context, this paper proposes a novel SI creation solution which formulates the SI fusion as a classification problem and employs a powerful classification algorithm, Support Vector Machine (SVM) [10], to efficiently select one out of several SI candidates. Since the content of video data is typically diverse, an online SVM learning approach is adopted where several SVM features are carefully defined and extracted. As assessed, the proposed SVM based SI fusion solution brings a major quality improvement with negligible additional complexity when compared to relevant SI creation solutions and can be easily integrated in the prior DSVC architectures [2, 9].

This paper is organized as follows: Section II presents the overall architecture of DSVC. Section III describes the proposed SI creation solution while Section IV presents and discusses the performance comparison for the proposed SI with relevant works. Finally, section V gives some conclusions and future works.

II. ARCHITECTURE OF DSVC

The adopted DSVC framework in this paper is the HEVC backward compatible distributed scalable video coding approach proposed in [9]. In order to understand the SI creation solution, this Section will briefly describe the general DSVC architecture.

A. Encoder architecture

DSVC encoder architecture with temporal and quality scalabilities can be illustrated in Figure 1.