Information Sciences 294 (2015) 439–455

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

Inter-embedding error-resilient mechanism in scalable video coding

Chia-Hung Yeh [a], Shu-Jhen Fan Jiang [a], Chih-Yang Lin [b,c,*], Min-Kuan Chang [d,e,*], Mei-Juan Chen [f]

[a] Department of Electrical Engineering, National Sun Yat-sen University, Kaohsiung 804, Taiwan
[b] Department of Computer Science and Information Engineering, Asia University, Taichung 413, Taiwan
[c] Dept. of Medical Research, China Medical University Hospital, China Medical University, Taichung, Taiwan
[d] Department of Electrical Engineering, National Chung Hsing University, Taichung 402, Taiwan
[e] Graduate Institute of Communication Engineering, National Chung Hsing University, Taichung 402, Taiwan
[f] Department of Electrical Engineering, National Dong Hwa University, Hualien 974, Taiwan

ARTICLE INFO

Article history:
Received 10 February 2014
Received in revised form 19 September 2014
Accepted 28 September 2014
Available online 12 October 2014

Keywords:
Scalable video coding
Error-resilient
Reversible data embedding

ABSTRACT

Scalable video coding (SVC), an extended version of H.264/AVC, is designed to transmit high-quality video bitstreams over heterogeneous networks. However, such video bitstreams are sensitive to transmission error, thereby severely degrading its quality. Inter-layer prediction in SVC causes errors to be propagated not only to subsequent frames but also to frames in the upper layers, when the errors have occurred in the lower layers. This paper presents an inter-embedding error-resilient scheme to reduce the distortion caused by the loss of inter-layer prediction information in SVC. The proposed algorithm exploits the reversible data embedding scheme to hide essential information of the lower layer without damaging the original data. Experimental results demonstrate that the proposed method provides a better PSNR performance than the frame copy by an average of 4.89 dB in a 2-layer SVC decoder, and an average of 4.54 dB in a 3-layer SVC decoder in the case of whole frame loss.

1. Introduction

Scalable video coding (SVC) [23], an extended version of H.264/AVC, is designed to support various devices with various bitrates and resolutions. One of the features of SVC is the layered coding of high-quality video that is composed of one or more subset videos. The complexity, reconstruction quality, and quantity of the SVC data in the subset bitstream are similar to those in H.264/AVC. To provide different video resolutions, bitrates, and video qualities, SVC uses several scalability modalities, including temporal, spatial, SNR/quality/fidelity scalabilities, or a combination of these modalities. In SVC, a multilayer video coding structure is utilized, which includes one base layer (BL) and multiple enhancement layers (ELs). It is called the pyramid coding scheme. The source video is the input of the EL and the BL. First, the source video is down-sampled for the BL. Then, the BL and the EL are coded independently. Only the information of mode, motion, and residual are extracted from the BL for transmitting to the EL based on the inter-layer prediction presented in Section 2.1.

* Corresponding authors at: Department of Computer Science and Information Engineering, Asia University, Taichung 413, Taiwan (C.-Y. Lin). Department of Electrical Engineering, National Chung Hsing University, Taichung 402, Taiwan (M.-K. Chang).

E-mail addresses: andrewlin@asia.edu.tw (C.-Y. Lin), minkuanc@nchu.edu.tw (M.-K. Chang).

http://dx.doi.org/10.1016/j.ins.2014.09.052
0020-0255/

However, because of the high compression ratio, video bitstreams are sensitive to transmission error, which severely degrades the video quality in SVC when the inter-layer prediction is applied; errors are propagated to following frames, and frames of the upper layers, when the errors occur in lower layers. Various mechanisms can be implemented to correct transmission errors and improve the quality of service.

Retransmission is one of these correction mechanisms, which involves retransmitting lost packets. However, such mechanisms produce information overhead that increases required bandwidth, reducing the transmission rate. Error concealment [4–6,8,15,21,28,36,38] is a decoder-based error correction tool, which provides a beneficial performance when the packet loss rate is low. Some researches adopt mode information together with motion vectors for error concealment purpose [27,30,32,33,40]. Zhao and Delp [38] proposed a slice interleaving method and used motion vectors of the lower layer to conceal errors. In one study [4], the frame copy (FC) and temporal direct motion vector generation (TD) are applied when the base layer is lost. If the enhancement layer is lost, motion and residual upsampling (BLSkip) and reconstruction base layer upsampling (RU) are used to conceal the errors. In another study [36], numerous decision rules are utilized to choose the TD or BLSkip method to conceal the error in the EL. Another study [21] proposed an error tracking model to estimate the distortion caused by error concealment and error propagation for the choice of BLSkip or RU. However, the PSNR decreases rapidly as the packet loss rate increases. In this situation, error resilience [2,7,9,11–14,17,18,20,22,25,26,35,41] outperforms error concealment.

Three papers [12,20,37] elucidated separate error-resilient methods for SVC based on multiple description coding (MDC) [2,22,26] and redundant coded picture [11,41], respectively. Zhao et al. [39] compared the performance between MDC and forward error correction based on different redundancy rates and average burst lengths. Zhang et al. [37] proposed an error-resilient method based on MDC containing two sequences: one is the original video sequence, and the other is the rotated video sequence. MD-SVC, which was developed by Mansour et al. [20], transmits information about motion and texture using MDC. The information about motion is the index of the neighboring block. The motion vector of the current block is resumed by the motion vector of the neighboring block if the current block is lost. The texture information is the residual data and the pixel value of intra-coded macroblocks (MBs). MD-SVC generates two descriptions for each EL. The first description includes the original motion vectors of even MBs, information about the motion of odd MBs, the original residual of odd MBs and information about the residual of MD-SVC of even MBs. The second description is the opposite of the first description. If only one description is successfully received, MD-SVC can provide satisfactory quality by using only one description. Jia et al. [12] developed redundant coded pictures, which contain downsampled residual information. When a transmission error occurs, the residual is resumed by interpolating the redundant coded picture.

Data embedding [1,7,9,13,14,17,18,25,35] is another scheme in error resilience. Kang and Leou [14] embedded the directions of the edge, modes, or motion vectors for error resilience in H.264/AVC. The odd/even data embedding approach [35] is one of the most popular methods used to hide information in video coding. However, most data embedding schemes change the original data and degrade the quality even when no packet is lost. The reversible data embedding scheme [25] can restore the original data without any distortion. Lie et al. [18] utilized this approach to embed required indices of wavelet coefficients for error concealment.

This study presents an error resilience method that is based on a reversible data embedding technique for SVC. The proposed scheme embeds the required information of the lower layer in its upper layer when an MB in the upper layer is inter-layer predicted. The preliminary results are published in [13]; this paper is more consistent in terms of theoretic analysis and experiments. The rest of this paper is organized as follows. Section 2 reviews inter-layer prediction in SVC and the reversible data embedding scheme. Section 3 describes the proposed algorithm in detail. Section 4 presents simulation results. Finally, Section 5 draws conclusions.

2. Background review

2.1. Inter-layer prediction in SVC

SVC further supports three new inter-layer prediction modes to improve the rate-distortion efficiency of the ELs for the EL encoding process. They are inter-layer motion prediction, inter-layer intra prediction, and inter-layer residual prediction. Since each layer has the same resolution in the quality scalability, the quality scalability can be regarded as a special case of the spatial scalability. Two inter-layer predictions cannot be employed in quality scalability: the upsampling operations and the inter-layer deblocking for intra-coded reference layer MB [23].

(1) *Inter-layer motion prediction:* In this prediction, the MB partition of the upper layer is obtained by upsampling the corresponding partition of the 8×8 block of the lower layer, and the predicted motion vector of the upper layer is also derived by scaling the corresponding motion vector of the lower layer. The search center of motion estimation is located at the predicted motion vector which is obtained by inter-layer motion prediction.
(2) *Inter-layer intra prediction:* When the BL is intra-coded, the MB of the EL can be predicted by upsampling the reconstructed sub-MB of the BL. The reconstructed sub-MB of the BL is upsampled by the four-tap filter for the luminance component and the bilinear filter for the chrominance component.
(3) *Inter-layer residual prediction:* This prediction is used for all inter coded blocks in the EL. The residual data of the corresponding 8×8 block of the BL is block-wise upsampled by a bilinear filter. The difference between the upsampling residual and the predicted one of the EL blocks needs to be coded.