Abstract

Based on the fact that high similarities exist between viewpoints of multi-view images, MV-HEVC achieves high encoding efficiency by performing conventional temporal direction prediction in a single viewpoint as well as inter-view prediction between viewpoints. In this paper, we propose to integrate SHVC and MVC (Multi-view Video Coding) to implement scalable multi-view video encoder using HEVC as a base layer. According to experimental results, it is verified that the BD-PSNR improvement reaches up to 1.5dB while reducing the BD-Bitrate by around 50~60%.

Keywords: Scalable Multi-view video coding, HEVC, SHVC, MV-HEVC

---

Design and Implementation of Scalable Multi-view Video Coding Based on Integration of SHVC and MVC

Tae-jun Jung and Kwang-deok Seo

Abstract

Based on the fact that high similarities exist between viewpoints of multi-view images, MV-HEVC achieves high encoding efficiency by performing conventional temporal direction prediction in a single viewpoint as well as inter-view prediction between viewpoints. In this paper, we propose to integrate SHVC and MVC (Multi-view Video Coding) to implement scalable multi-view video encoder using HEVC as a base layer. According to experimental results, it is verified that the BD-PSNR improvement reaches up to 1.5dB while reducing the BD-Bitrate by around 50~60%.

Keywords: Scalable Multi-view video coding, HEVC, SHVC, MV-HEVC
I. Introduction

To provide optimum quality multi-view video service under heterogeneous network environments and mobile equipment, it is needed to encode multi-view video in a layered format based on scalability [1], [2]. So it is important to develop a scalable multi-view video coding structure which combines multi-view video coding technology, MV-HEVC, and layered video coding technology, SHVC [1]. But there must be an update in the reference picture management when we combine the two separate codecs to develop a scalable multi-view encoder. Reference picture is generally managed by reference picture list. To perform picture-to-picture prediction efficiently in the process of video encoding and decoding, many reference pictures should be stored in the memory. And the DPB (Decoded Picture Buffer) also needs to be changed to accommodate the developed scalable multi-view coding structure [3].

In this paper, we propose an efficient design and implementation of scalable multi-view video coding based on integration of SHVC and MVC. The proposed design mainly consists of three parts: 1) reference picture reordering part which reorganizes the sequence of the reference pictures, 2) marking process part which stores and dispatches reference pictures to and from the DPB memory, and 3) reference picture list reconstruction part which composes the two conventional reference picture lists, each one used for SHVC and MV-HEVC. With the proposed reference picture list design, the predictive coding of the developed scalable multi-view video coding can be performed efficiently in terms of advanced prediction accuracy in motion estimation, systematic DPB management, and improved coding efficiency.

II. Design Of Reference Picture List and Prediction Mechanism

Fig. 1 shows the system architecture of the proposed scalable multi-view video coding based on integration of SHVC and MVC for dual-views video, where major coding structures and mechanisms of SHVC and MV-HEVC are integrated. Each view is basically processed by applying the layered coding mechanism of SHVC to generate the base and enhancement layers in a scalable way. To achieve inter-view prediction effect, the reconstructed picture of view0 is used as an additional reference frame for coding view1.

![Diagram](Image)

그림 1. SHVC 및 MVC 겸용 기반의 세련된 스펙트럼리볼루션 이상동 비디오 부호화 구조

Fig. 1. System architecture of the proposed scalable multi-view video coding based on integration of SHVC and MVC

Fig. 2 shows the overall architecture of the designed reference picture list for the proposed scalable multi-view video coding based on integration of SHVC and MVC employing inter-view prediction. In this architecture, predictive coding efficiency can be improved by using enhancement layer picture of view0 (B0_2) as an additional reference picture for coding the picture of view1 (B1_2).
Fig. 2. Overall architecture of the designed reference picture list supporting inter-view prediction

Fig. 3 shows prediction structures between layers. Each of them is denoted as follows: 1) simulcast case: all layers are coded independently with HM (HEVC test model) software, 2) MV-HEVC simulcast case: layers are coded with HTM (MV-HEVC reference software) where D1 and D0 are coded independently, 3) SHVC simulcast case: layers with spatial dependency are coded with SHM (scalable HEVC reference software) where V0 and V1 are coded independently. 4) MV-HEVC+SHVC case: in addition to inter-view prediction predicted from D0 at the same view using inter-layer prediction, D1 is predicted from D0 at the same view using inter-layer prediction. In this implementation, we adopted the prediction mechanism of 4) MV-HEVC+SHVC case.

III. Experimental Results

In order to evaluate the compression efficiency and video quality after applying the proposed mechanism in actual coding process, we developed a scalable multi-view video coding system which integrates SHM-10.0 [4] and HTM-13.1 [5] together. Test condition for the performance evaluation was set to GOP_size=8, Intra_period=16, and frame_rate=25 frames per second. We used Poznan Hall HD sequences which are composed of dual views, one for the left-view and the other for the right-view. The maximum number of reference frames for motion estimation was set to 6. We compare the performance in terms of compression efficiency and average PSNR between the bitstreams generated by the conventional SVC+MVC predictive coding mechanism and by the proposed predictive coding mechanism integrating SHVC and MVC. The test is performed using four different values of QP (quantization parameter) values such as 27, 30, 33, and 35.

Table 1 shows the test results for coding the left-view video of the Poznan Hall sequences with HD resolution. “SVC prediction” in Table 1 denotes the results of coding for the left-view video by using MVC+SVC predictive coding architecture. “Proposed prediction” denotes the results obtained by using the proposed predictive coding architecture. The experimental results in Table 1 show that we could achieve...
1.4697 dB increase in BD-PSNR and 57.5634% decrease in BD-Bitrate at the same time by employing the proposed prediction mechanism. Fig. 4 shows the comparison of PSNR performance by plotting RD (rate-distortion) curve for the Poznan Hall test sequences.

### IV. Conclusions

In this paper, we proposed an efficient design and implementation of scalable multi-view video coding based on integrating SHVC and MVC mechanism. Experimental results showed the superiority of the proposed scalable multi-view video coding compared to the conventional implementation in terms of compression efficiency and video quality.

### 참고 문헌 (References)


