Available online at www.elixirpublishers.com (Elixir International Journal)

Computer Science and Engineering

Elixir Comp. Sci. & Engg. 42 (2012) 6332-6334

ISSN: 2229-712X

An inter-view prediction technique using motion skip for multi-view video coding

Krishna Reddy K.R[1] and G.Mahadevan[2]

[1]PRIST University, Thanjavur, Thamilnadu, India,

[2]Department of Computer Science and Engineering, AMC College of Engineering, Bangalore –83, Karnataka, India.

ARTICLE INFO

Article history:

Received: 11 October 2011;

Received in revised form:

5 January 2012;

Accepted: 18 January 2012;

Keywords

Multiview Video Coding,

Single motion compensation loop Decoding,

Motion Skip,

H.264/AVC.

ABSTRACT

Multiview video coding is an extension of H264/AVC. When an Multiview video coding bitstream is decoded, some views (named target views) are to be is played; some other views (named dependent views) may not be displayed but are needed for inter-view prediction of the target views. The original Multiview video coding design requires pictures of the dependent views to be fully decoded and stored. This entails both high decoding complexity and high memory consumption for the pictures in the views which are not intended for display, particularly when the number of dependent views is large. In this paper, a single motion compensation loop decoding scheme is introduced to address these disadvantages. The proposed scheme requires only partial decoding of pictures in dependent views and thus significantly reduces decoding complexity and memory consumption. The proposed method is based on the motion skip, wherein inter-view motion and coding mode prediction is exploited. Simulation results shows that the proposed scheme provides a substantial reduction of complexity and memory size, at the expense of only a minor compression efficiency loss, compared with multiple motion compensation loop decoding schemes for Multiview video coding.

Introduction

Multiview video technologies have gained increasing interest recently. In multiview applications, the original video content is a group of video sequences captured by multiple cameras at the same time from the same scene through different viewpoints. The joint video team (JVT) is developing a multiview video coding technique [1], which will become an extension of H.264/AVC. It takes advantage of the inter-view correlation to improve coding efficiency and also provides bit-rate and view scalability, error robustness, and control of decoding complexity [2].

Among typical Multiview video applications, such as free viewpoint video [3], 3D TV [4], and immersive teleconferencing, there are cases that display only a subset of the encoded views. In this paper, the views that are displayed are referred to as the target views or output views, while the remaining views are referred to as the dependent views. Target views are related to dependent views through inter-view prediction. For example, in free-viewpoint video, only one view needs to be displayed at a certain time instance and that view is the target view.

In the latest joint draft (JD) of Multiview video coding [2], inter-view prediction is realized by utilizing pictures from other views as reference pictures for motion compensation. This interview prediction is referred to as inter-view sample prediction. In addition to the JD, JVT also maintains a joint multiview video model (JMVM) [5] for Multiview video coding. JMVM documents some additional tools that were shown to be potentially useful but not mature enough to be included into the JD. In the JMVM, there is a tool called motion skip, which enables motion prediction between views.

Single motion compensation loop decoding [6] is supported in the scalable extension of H.264/AVC, also known as Scalable video coding [7]. The basic idea of single motion compensation loop decoding in Scalable video coding is as follows. To decode a target layer that depends on a number of lower layers, only the target layer itself needs to be fully decoded, while for the lower layers only parsing and decoding of Intra macroblocks (MBs) are needed. Essentially, Single motion compensation loop decoding in Scalable video coding requires motion compensation only at the target layer. Consequently, Single motion compensation loop decoding provides a substantial complexity reduction. Furthermore, since the lower layers do not need motion compensation and no sample values need to be stored in the decoded picture buffer (DPB), the decoder memory requirement is greatly reduced compared to conventional multiple motion compensation loop decoding , where motion compensation and full decoding is needed in every layer.

In this paper, single motion compensation loop decoding is applied to Multiview video coding similarly as it is applied to Scalable video coding. In the proposed method for Multiview video coding, only the target views are fully decoded, while the non-anchor pictures of the dependent views need only to be parsed to obtain information required for inter-view prediction.

This paper is organized as follows. In Section 2, we explain how the motion prediction can be realized in Multiview video coding. In Section 3, illustrations are given for the proposed single motion compensation loop decoding. Simulation results are provided in Section 4 and Section 5 concludes the paper.

Tele:

E-mail addresses: krkrishnareddy@rediffmail.com

Motion Prediction in Scalble video and Multiview video

View dependency

In the draft Multiview video coding specification, an access unit consists of pictures of all the views pertaining to a certain display or output time. An anchor picture is a picture in a view that can be correctly decoded without the decoding of any earlier access unit in decoding order (i.e. bitstream order). Anchor pictures can serve as random access points. Other pictures are non-anchor pictures. View dependency is specified in the sequence parameter set (SPS) Multiview video coding extension, separately for anchor pictures and non-anchor pictures.

A typical prediction structure of Multiview video coding is shown in Fig. 1, wherein anchor pictures are those of the time instances T0 and T8. Pictures within each view form a hierarchical temporal prediction structure. Prediction across views is also enabled, but is constrained in the same time instance.

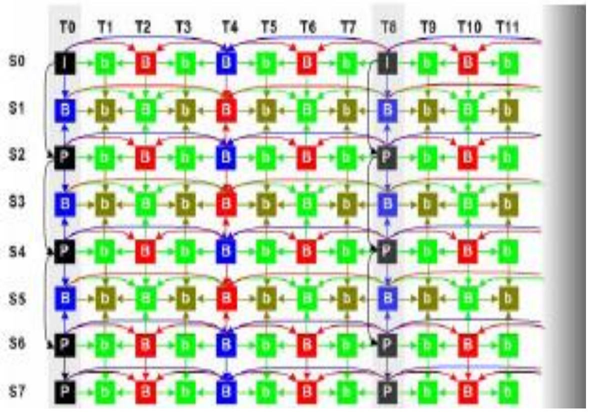

Fig. 1. Typical Multiview prediction structure with 8 views and GOP length

For a certain view, dependent pictures in other views are referred to as inter-view pictures. According to Multiview video coding JD, the texture (i.e. reconstructed sample values) of those pictures is required since inter-view pictures are used as reference pictures for motion compensation.

For bi-predictive (B) pictures, dependent views can be signaled in two prediction directions, corresponding to the reference picture list (we use list for simplicity) 0, and list 1.The views corresponding to list 0 (or list 1) are named forward (or backward) dependent views. For example, in Fig. 1, view 0 is the forward dependent view of view 1,while view 2 is the backward dependent view of view 1. A view that has both forward dependent views and backward dependent views is called a B-view. If a view has only forward dependent views, it is called a P-view.

Inter-layer motion prediction in Scalable video

In Scalable video coding, pictures in different layers of the same access unit coincide in time and thus the corresponding motion vectors have strong correlation.

So, motion vectors of an MB are derived from the co-located MB(s) in the base layer. In spatial scalability, wherein resolutions of layers are different, base layer motion vectors are scaled. In CGS (Coarse Granularity Scalability), wherein layers have the same resolution, the derived motion vectors are the same as those of the co-located MB at the base layer. In Scalable video coding, the derived motion vectors can be further refined, and the refinement is coded into the bitstream.

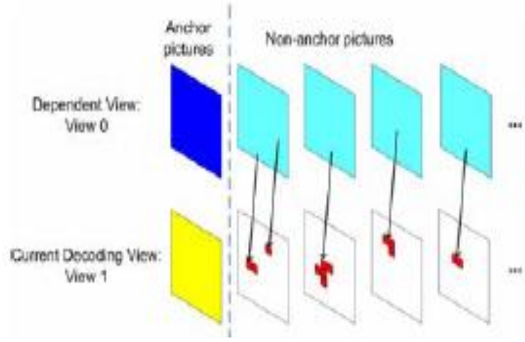

Fig. 2. Motion skip using disparity motion vector.

Motion skip in JMVM

In Multiview video coding, different views are captured by different cameras. This makes inter-view motion prediction in Multiview video coding harder and less efficient compared to inter-layer prediction of Scalable video coding. As a coding tool in JMVM, motion skip predicts modes and motion vectors from the inter-view pictures and it applies to non-anchor pictures only. During encoding, a global disparity motion vector (GDMV) is estimated when encoding an anchor picture, and then GDMVs for nonanchor pictures are derived so that the GDMVs for a nonanchor picture is a weighted average from the GDMVs of the two neighboring anchor pictures.

A GDMV is of 16-pel accuracy, i.e., for any MB in the current picture (i.e. the picture being encoded or decoded), the corresponding region shifted in an inter-view picture according to the GDMV covers exactly one MB in the inter-view picture. Normally, the first forward dependent view is considered for motion skip.

However, if the corresponding MB in this picture is Intra coded, then the other candidate, which can be the first backward dependent view, if present, is considered. For an MB, if the motion skip mode is selected, motion vectors are derived and only residual between the original signal and the motion compensated signal are transmitted.

An example of motion skip is shown in Fig. 2, wherein view 0 is the dependent view and view 1 is the target view. With the disparity motion, when decoding MBs in view 1, the corresponding MBs in view 0 are located and their modes and motion vectors are reused as the MB modes and motion vectors for the MBs in view 1.

Unlike inter-view sample prediction, the use of which entails Multiple motion compensation loop decoding since it requires motion compensation for inter-view pictures, the use of motion skip does not require motion compensation of inter-view pictures. Thus, motion skip can be used for proposed single motion compensation loop decoding.

Proposed decoding scheme for Multiview video coding

To achieve single motion compensation loop decoding for coding of non-anchor pictures, only motion skip is applied for inter-view prediction. Inter-view sample prediction is only used for coding of anchor pictures. A flow chart for the proposed method at the decoder is shown in Fig. 3. When decoding an anchor picture, inter-view sample prediction can be used and the picture is fully decoded and stored into the decoded picture buffer. When decoding a non anchor picture, inter prediction and motion skip can be used, if this non-anchor picture belongs to a target view, it is motion compensated, reconstructed and stored in the decoded picture buffer, otherwise, it is only parsed and only its coding modes and motion field are constructed.

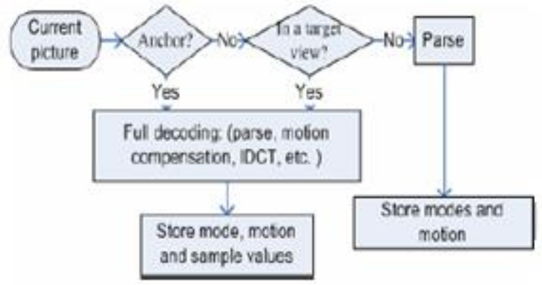

Fig. 3.Proposed scheme in multiview video coding

An example of decoding complexity reduction and memory saving in the free-viewpoint video case is given for the prediction structure shown in Fig. 1.. It is assumed that view 5 is the only target view. When the multiple motion compensation loop decoding scheme is in use, all the pictures in views 0, 2, 4, 5 are to be fully decoded and stored in the decoded picture buffer in order to successfully decode view 5. However, if the proposed scheme is in use and a non-anchor target picture is decoded, the dependent pictures in the same access unit need only to be parsed for generation of the coding modes and motion vectors, while their sample values do not need to be constructed.

Simulation Results

The proposed method was implemented based on the reference software of JMVM, and compared to the MVC Multi motion compensation loop decoding. The standard Sequences defined in the MVC common test condition [8], namely *Akko&Kayo, exit, ballroom*, and *rena* were tested.

In Table 1, the performances of proposed method and conventional decoding method with motion skip are compared. Results are generated using the Bjontegaard measurement [9]based on the bitrates and average luma peak signalto-noise (PSNR) values of the four test points corresponding to QP values 22, 27,32, and 37. Note that for proposed method, we enabled motion skip for non-anchor pictures in all the views, while for conventional method, inter-view prediction is enabled only for the B-views. Even with the additional prediction dependency, proposed method is with comparable complexity and memory consumption in the P views and with much less complexity and memory computation in the B-views. So the total computational complexity and buffer required for the motion skip in proposed method are still lower than Multiple motion compensation loop decoding.

Conclusions

In this paper, a single motion compensation loop decoding scheme is proposed for multiview video coding. In this scheme, anchor pictures are coded with inter-view sample prediction and intra prediction while the non-anchor pictures are coded with motion skip, inter prediction, and intra prediction in the joint multiview video model (JMVM). The simulation results indicated that the proposed scheme entails a minor compression efficiency loss (upto 0.13dB approx.)compared to multiple motion compensation loop decoding scheme but reduces decoder complexity and memory usage remarkably.

References

[1] "Joint Draft 5.0 on Multiview Video Coding," *JVT-Y209*, Shenzhen, China, Oct. 2007.
[2] ISO/IEC JTC1/SC29/WG11, "Requirements on Multi-view Video Coding v.5," *N7539,* Nice, France, Oct. 2005.
[3] A. Smolic, and P. Kauff, "Interactive 3D Video Representation and Coding Technologies", *Proc. IEEE, Special Issue on Advances in Video Coding and Delivery,* Jan. 2005.
[4] A. Vetro, W. Matusik, H. Pfister, J. Xin, "Coding Approaches for End-to-End 3D TV Systems," *Picture Coding Symposium*, 2004.
[5] "Joint Multiview Video Model (JMVM) 6.0," *JVT-Y207*,Shenzhen, China, Oct. 2007.
[6] "video decoding in spatial scalability using inter layer Prediction", Krishnareddy K.R, Naghabhushna, Aravind T.V.,Dr.G.mahadevan, india, July - Aug 2010.
[7] "Joint draft 11 of SVC amendment," *JVT-X201*, Geneva, Switzerland, Jun.-Jul. 2007.
[8] "Common Test Conditions for Multiview Video Coding," *JVT-T207*, Klagenfurt, Austria, Jul. 2006.
[9] G. Bjontegaard, "Calculation of average PSNR differences between RD-curves," *VCEG-M33*, Mar.2001.

Table 1. Comparison of Single (proposed) and Multiple (conventional) motion compensation loop decoding

Sequence	Proposed method	Conventional method
	Average PSNR-Y(dB)	Average PSNR-Y(dB)
Akko &Kayo	35.65	35.682
Exit	37.44	37.42
Ballroom	35.7875	35.9165
Rena	40.17	40.011