Signal Processing: *Image Communication* 29 (2014) 667–677

Contents lists available at ScienceDirect

Signal Processing: *Image Communication*

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

Efficient multi-view video coding using inter-view information

Xin-Xian Huang [a], Mei-Juan Chen [a], Chia-Hung Yeh [b,*], Hao-Wen Chi [a], Chia-Yen Chen [c]

[a] Department of Electrical Engineering, National Dong Hwa University, Hualien 974, Taiwan, ROC
[b] Department of Electrical Engineering, National Sun Yat-sen University, Kaohsiung 804, Taiwan, ROC
[c] Department of Computer Science and Information Engineering, National Kaohsiung University, Kaohsiung 811, Taiwan, ROC

ARTICLE INFO

Article history:
Received 4 October 2013
Received in revised form
12 February 2014
Accepted 20 March 2014
Available online 13 April 2014

Keywords:
Multi-view video coding
Disparity
Motion estimation
Mode decision
Rate-distortion cost
Search range

ABSTRACT

Multi-view video coding (MVC) has been extended from H.264/AVC to improve the coding efficiency of multi-view video. This paper proposes a fast mode decision algorithm which can make an early decision on the correct mode partition to solve the issue of the enormous computational complexity. The best modes of the reference views are utilized to determine the complexity of the macroblock (MB) in the current view, the mode candidates needed to be calculated can then be obtained according to the complexity. If the complexity is low or medium, the search range can be reduced. The threshold of the rate-distortion cost for the current MB is calculated using the co-located and neighboring MBs in previously coded view and is utilized as the criterion for early termination. The motion vector difference in the reference view is applied to dynamically adjust the search range in the current MB. Experimental results prove that the proposed algorithm achieves a time saving of 81.05% for a fast TZ search and 87.85% for full search, and still maintains quality performance and bitrate.

1. Introduction

Multi-view video coding (MVC) is used to process 3D video and is a practical technology that utilizes different views to present the desired information. Various coding frameworks of MVC based on H.264/AVC have been developed over the past few years. However, a multi-view video contains a large amount of data, which requires considerable coding time and is a critical issue in time-consumption. Therefore, the Joint Video Team (JVT) has developed the joint multi-view video coding (JMVC) to improve coding efficiency [1]. Fig. 1 shows the reference frames from the inter-view and the time domain for achieving high coding efficiency [1].

The coding architecture of the frames in a group of pictures (GOP) is a hierarchical B picture (HBP) structure, and the coding architecture of the views is divided into two parts: the reference directions with only temporal domain in the main views (S0, S2, S4 and S6), and the reference directions with the temporal domain as well the inter-view in the auxiliary views (S1, S3, S5 and S7). As a result, there are two kinds of predictive directions in the auxiliary views. Lowering the computational complexity in MVC has become an interesting topic to many researchers. In recent years, many studies have proposed many fast mode decision algorithms, which are classified into four types as follows. The first type uses early termination determined by a threshold to save time [2–6]. The second type determines the reference direction for motion estimation [7,8] or

* Corresponding author. Tel.: +886 7 5252000x4112; fax: +886 7 5254199.
E-mail address: yeh@mail.ee.nsysu.edu.tw (C.-H. Yeh).

http://dx.doi.org/10.1016/j.image.2014.03.005
0923-5965/

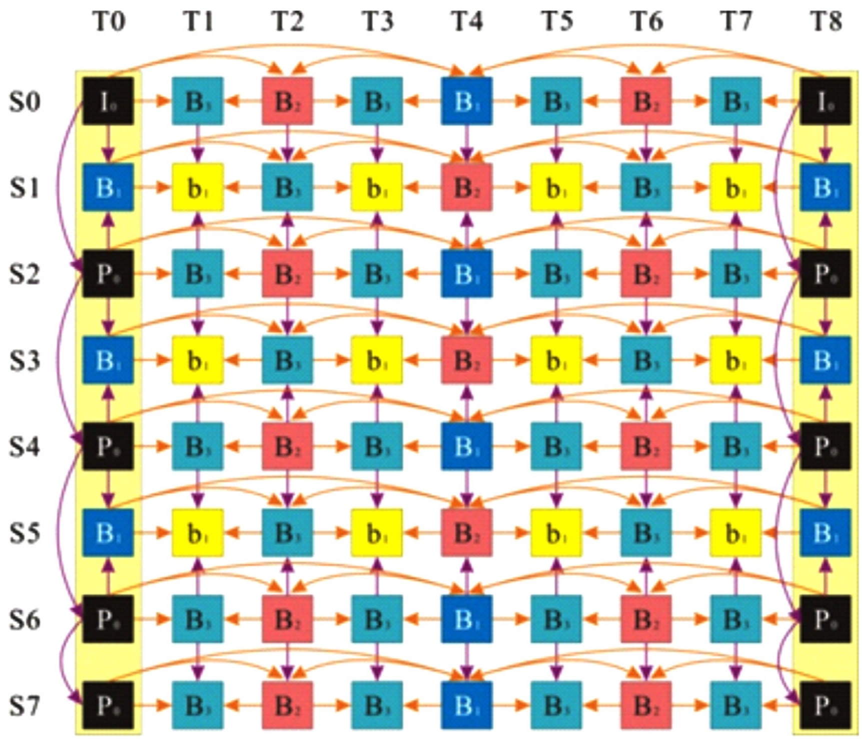

Fig. 1. MVC coding architecture [1].

speeds up motion and disparity search [9]. The third type performs similarity analysis between the current MB and the co-located MB in the neighboring view [10–17]. The fourth type jointly uses the search range, reference direction and the MB complexity [18–21]. However, the above mentioned methods still require improvements with respect to time and coding quality. Methods with higher reduction in time often lead to poor image quality in the coded video, whereas methods having better PSNR and compression rate often require more coding time.

This paper is organized as follows. The analysis of various modes and the proposed algorithm are presented in Section 2, extensive simulation results are provided in Section 3 to demonstrate the proposed algorithm is able to provide faster coding with good image quality, and finally, the conclusion is given in Section 4.

2. Proposed fast mode decision algorithm

The proportion of the various modes of each test video sequence in the auxiliary view is discussed first. The sequences in use are *Ballroom*, *Exit*, *Jungle*, *Race1*, *Uli* and *Vassar* video sequences. The coding parameters are as followings, quantization parameter (QP) of 24 to 36, two reference frames, search range of 64, and 250 encoding frames. The statistics of the mode distributions in the video sequences with respect to QP of 24, 28, 32 and 36 are exhibited in Fig. 2(a) to (d), respectively. From Fig. 2(a) and (b), it can be seen that with QP of 24 and 28, Direct mode occupies the majority of the modes, followed by Mode 16 × 16. This phenomenon is even more apparent in Fig. 2(c) and (d), with QP of 32 and 36, where the Direct mode occupies more than 55% and Mode 16 × 16 occupies over 10% in most sequences, while the other coding modes, chosen from the best coding modes, are all much less; however, these other modes all require much more computation complexity than Direct mode and Mode 16 × 16.

2.1. Complexity classification of the macroblock

In Fig. 3, the four prediction directions for a MB are presented. The four prediction directions include two forward reference directions, T_{list0} and V_{list0}, and two backward reference directions, T_{list1} and V_{list1}, for the complexity calculation for the MB. T represents the temporal reference direction and V represents the inter-view reference direction. The complexity for the current MB can be obtained from the inter-view and temporal domain modes. The complexity detection for the four reference directions is based on Eq. (1). D_n indicates the reference direction. In Eq. (1), the distance for reference frame, S, is determined first. In Fig. 4(a) and (b), when the current view is $S1$, $S0$ and $S2$ are considered as previously coded views and will be used to calculate complexity for the current MB. If S is equal to 1, all four reference directions, as shown in Fig. 4(a), should be considered because the probability of Direct mode is very high in a HBP structure. Therefore, the four reference directions are considered to avoid an erroneous judgment. If S is not equal to 1, the reference direction which has the minimal rate-distortion cost (RD cost) from the co-located MBs among these four reference directions would be used as the final reference direction. The modes of the co-located MB and the eight neighboring MBs are