J. Vis. Commun. Image R. 24 (2013) 686–699

Contents lists available at SciVerse ScienceDirect

J. Vis. Commun. Image R.

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

Coding statistics based fast mode decision for multi-view video coding

Chia-Chi Chan, Chih-Wei Tang *

Visual Communications Laboratory, Department of Communication Engineering, National Central University, Jhongli 32001, Taiwan

ARTICLE INFO

Article history:
Available online 18 January 2012

Keywords:
Multi-view video coding
Fast mode decision
RD cost
Motion cost
Motion vector difference
DIRECT mode
Inter mode
Intra mode

ABSTRACT

Reduction of high computational complexity of multi-view video coding (MVC) is necessary for realization in consumer electronics. Since mode decision is one of the key computational bottlenecks of multi-view video encoders, this paper proposes a coding statistics based fast mode decision algorithm. First of all, a rate–distortion cost based fast DIRECT mode decision algorithm early terminates the mode decision process if possible. Next, the candidates for Inter modes are reduced by taking the advantage of the correlation between an optimal mode and motion cost. The proper thresholds to reduce the candidates for the above two fast algorithms can be easily derived from exponential functions at run time. Finally, motion vector difference based motion characteristics is referred to further speed up the mode decision process of Inter modes. The experimental results show that the proposed scheme reduces up to 70.82% of encoding time with negligible degradation of RD performance.

1. Introduction

Multi-view video coding is essential to the success of 3D applications such as free viewpoint television (FTV) [1] and three-dimensional television (3DTV) [2,3]. Joint Video Team (JVT) has finalized the multi-view video coding standard extended from the single view video coding, H.264/AVC, in early 2008 [4]. The encoder complexity distribution of AVC has been analyzed in [5]. Thus, fast mode decision is important and examples of fast mode decisions for H.264/AVC can be found in [6,7]. Different from single view video coding, MVC provides both temporal and inter-view predictions to reduce data redundancy. For each MB, the best motion vector, disparity vector, and block mode are decided by rate–distortion optimization (RDO) [8]. Although the optimal RD performance can be achieved with full search, motion and disparity estimations along with mode decision are the most time-consuming modules of a MVC encoder. Our analyses show that their time complexity occupies about 97% of the total MVC encoder complexity. Therefore, significant time saving can be achieved with fast mode decision and motion estimation.

In recent years, several fast encoding algorithms have been proposed for MVC. In [9], a two-stage neural classifier is adopted to decide the candidate modes based on block partitions and mode distribution analysis of stereo video coding. In [10], four fast mode decision algorithms are proposed to reduce the computational complexity of motion estimation and disparity estimation for MVC. The candidates are reduced according to the coding information of the corresponding MB and neighboring MBs in the neighboring view. In [11], fast mode decision analyzes the average RD cost and mode proportion of mode category in non-anchor frames of the base view. Then, it reduces modes by comparing RD cost with an adaptive threshold that varies with QP and average RD cost. In non-anchor frames of the other views, fast mode decision uses the mode correlation between neighboring views for early termination of coding process. Since DIRECT mode occupies the dominant proportion of optimal modes and requires the least encoding time, a fast DIRECT mode decision algorithm can effectively boost the mode decision process. In [12], the early DIRECT mode decision is achieved based on the statistical characteristic of RD cost, variance property of DIRECT mode, and neighboring modes along both temporal and inter-view prediction directions. In [13], DIRECT mode is determined to be the optimal mode or not by comparing the RD cost of DIRECT mode with an adaptive threshold, derived from the coding information and weighting factors of the neighboring blocks in the previously coded view and the co-located block in the reference frame. In addition to fast mode decision algorithm for MVC, Micallef et al. [14] propose a fast motion and disparity estimation algorithm by using the depth information of the video sequence and multi-view geometry to reduce the motion and disparity vector search windows.

Since the RD cost of mode decision is highly related to the RD cost of motion estimation (i.e., motion cost), in [15], a motion cost based early termination algorithm is designed for fast decision of Inter modes in H.264/AVC. According to the conditional probability of each mode, the corresponding threshold of early termination is derived for each MB. The algorithm is further extended in [16] to enable the fast mode selection of Inter modes and sub-MB modes in H.264/AVC, where the correlation of mode decision and motion cost is explored. Inspired by Ma et al. [16], for

* Corresponding author.
E-mail address: cwtang@ce.ncu.edu.tw (C.-W. Tang).

1047-3203/$ - see front matter © 2012 Elsevier Inc. All rights reserved.
doi:10.1016/j.jvcir.2012.01.004

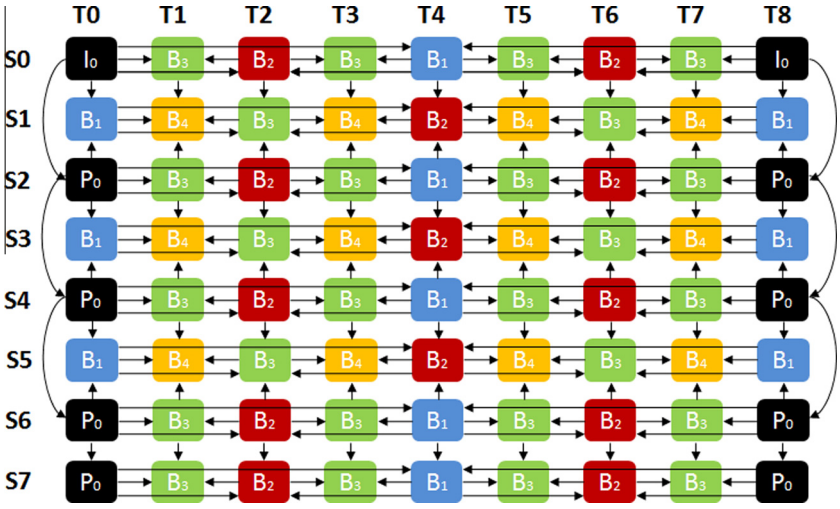

Fig. 1. Our adopted prediction structure for multi-view video coding.

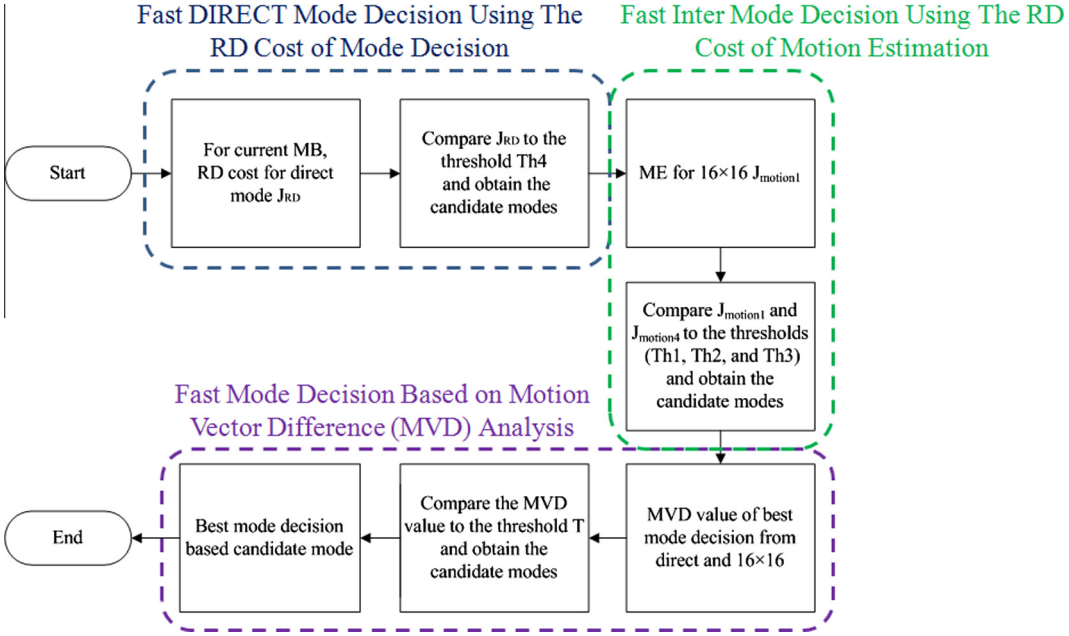

Fig. 2. The flowchart of proposed coding statistics based fast mode decision for multi-view video coding.

each multi-view video, Chan et al. [17] decides a set of dynamic thresholds for each view individually using the on-line statistical analysis of motion and disparity costs of the first GOP. The set of candidates for Inter mode decision in the rest of frames is determined by these thresholds. Fast mode decision can be also achieved by classifying the image content using the coding information. For example, RD costs of Intra4 $\times$ 4 and Intra16 $\times$ 16 from neighboring views can be used to classify a textural region into a low region, medium region, or high region. Inter8 $\times$ 8 mode will not be considered if the current MB is a low region or median region [18]. In [19], according to the motion vectors of the corresponding MB and neighboring MBs in the previously coded view, the classes of motion homogeneity of blocks are divided into homogeneous motion and complex motion and thus the candidates of Inter modes are selected.

Motivated by the previous work, this paper proposes a coding statistics based fast mode decision algorithm for MVC, where candidates including DIRECT, Inter, and Intra modes are considered to be reduced during the fast mode decision process. The inter-block motion information and RD costs of motion estimation and mode decision are referred since all of them have high correlation with mode decision. The proposed algorithm is stated as follows. Initially, a novel fast DIRECT mode decision algorithm, using the RD cost of DIRECT mode of the current MB and the histogram of RD cost of Inter16 $\times$ 16 mode obtained off-line, early terminates the mode decision process if possible. Next, candidates of Inter modes can be reduced using motion costs of MB and sub-MB modes. The adaptive thresholds for fast mode decision process can be easily derived from exponential functions given QP, where the parameters of exponential functions are analyzed off-line. Finally, candidates of Inter modes can be further filtered out based on the temporal coding information and the characteristic of motion vector difference that describes the inter-block motion variation. The rest of the paper is organized as follows. Section 2 describes the overall fast mode decision algorithm that combines the algorithms proposed in Sections 3 and 4. In Section 3, we describe the pro-