Adaptive Search Range Determination for Geometry based Disparity and Motion Estimation of MVC

Narasak Boonthep, Werapon Chiracharit, Kosin Chamnongthai
Department of Electronic and Telecommunication Engineering
Faculty of Engineering
King Mongkut's University of Technology Thonburi
126 Pracha Uthit Rd., Bang Mod, Thung Khru, Bangkok 10140, Thailand
E-mail: narasakb@gmail.com, werapon.chi@kmutt.ac.th, kosin.cha@kmutt.ac.th
Tel: +66 2470 9064, Fax: +66 2470 9070

Abstract—Multi-view video consists of a set of multiple video sequences from multiple viewpoints or view directions in the same scene. It contains extremely a large amount of data and some extra information to be stored or transmitted to the user and computational complexity of processing. In order to achieve a better coding performance in multi-view video coding (MVC), the search range for both motion estimation (ME) and disparity estimation (DE) are set twice as large as necessary for ME alone, which brings great computational complexity to the encoder. Search range (SR) of motion search constraints current block to search the matched block within a predefined area in the reference picture. This research proposed an algorithm to solve the redundancy reduction problem to get fast disparity and motion estimation for reducing complexity. We use geometry and adaptive search range by adjust macroblock (MB) type that suitable with window SR. Experimental results show that the proposed method can reduce the encoding time, so that the encoder complexity can be reduced but achieves almost the same compress quality. The encoding time reduce up more than 70%.

I. INTRODUCTION

Nowadays, the world’s technologies are developing and changing all the time. In the past, users are limited to see only 2D so they can access to the image only one side but 3D technology allow them to access the image with freedom view so it seem more realistic video to users. The viewer can interactively choose viewpoint in 3D space to observe the action of a dynamic real-world. Visual Communication Technology is one of the most important technologies that have been improved in many fields such as 3D image processing, Holography, Multi-View Video- MVV. Especially, Multi-view video have been applied into many applications like Free Viewpoint Video (FVV), Free Viewpoint Television (FVT), and Video-teleconferencing and 3DTV.

Multi-view video (MVV) [2, 3, 4, 5, 6] consists of a set of video sequences captured the same scene simultaneously from cameras but at difference view directions as show in Fig. 1. It contains a huge of data and a high degree of correlation within each view as well as among the views exploiting as inter-view statistical dependencies between the camera views in addition to temporal statistical dependencies within each sequence. A full-scale multi-view video system allows interaction with scene immersion, but also demands extensive storage and transmission bandwidth due to the massive data volume. A multi-view video contain multiple views of the same scene and some extra information capturing the correlation among these views, efficient compression of MVV is a key enabling factor for 3DTV and FTV applications. Multi-view Video Coding (MVC) must exploit the redundancies among the views. However, the exploitation of extra redundancies incurs extensive computation, counteracting the benefit gained from coding efficiency.

The coding efficiency is also consider other constraints and requirements such as computational and memory complexity, efficient random access in time for certain views, delay, buffer management, as well as parallel processing. Furthermore, specific coding modes for MVC bring new focus to the prediction structure. It has been shown that such dedicated MVC methods outperform the independent coding of MVV significantly while employing the same core codec in this approach is aimed to improve compressing efficiency. Many 3DTV systems are based on scenarios, where a 3-D scene is captured by a number of N cameras [7][8]. The simplest case is classical stereo video with two videos. More advanced systems apply 8, 16, and more cameras. Some systems additionally apply per sample depth data that can also