Improving 3D-TV View Synthesis Using Motion Compensated Temporal Interpolation

Minh Le Dinh [a,*], Long Vuong Tung [a], Xiem Hoang Van [a], Duong Dinh Trieu [a], Tung Pham Thanh [b], Ha Le Thanh [a]

[a]University of Engineering and Technology, Vietnam National University, Hanoi, 144 Xuan Thuy, Cau Giay, Hanoi, Vietnam.

[b]University of Fire Fighting and Prevention, 243 Khuat Duy Tien, Thanh Xuan, Hanoi, Vietnam.

***Abstract*—Nowadays, the development of three-dimension (3D) video applications such as three-dimensional television (3D-TV) and free-viewpoint television (FTV) has greatly increased human experiences. View synthesis method like depth-image-based-rendering (DIBR), plays a significant role in 3D content creation, 3D transmission, and has been integrated into video coding standards such as 3D-High efficiency video coding (3D-HEVC). However, the current DIBR method employs only the disparity correlation between views to create a so-called synthesized view; thus, unable to take full advantages of available synthesized information. In this paper, we propose a novel view synthesis method which takes advantages of not only the disparity correlation but also the temporal correlation between views. In the proposed method, an effective motion compensation based frame interpolation is employed to generate a temporal prediction view which is then combined with the DIBR rendered view to obtain the final synthesized view. Experimental results show that the proposed method can achieve the synthesized view with significantly outperforming other conventional techniques in terms of both peak signal-to noise ratio (PSNR) and subjective visual quality.**

***Index Terms*—View synthesis, Depth-image-based rendering, VSRS-1D Fast.**

I. Introduction

Three-dimensional television (3DTV) and its next generation - Free viewpoint television (FTV) are technology promise to bring incredible 3D experience, viewers not only watch 3D content but also immerse themselves there. FTV is regarded as the ultimate 3DTV because it enables realistic viewing and free navigation of 3D scenes [1].

FTV has been developed to deliver an infinite number of views since the viewpoint can be placed anywhere. It is impossible to capture and transmit all views of the scene so, in the typical scenario of FTV, there is a view synthesis module to create new views from given captured views (also known as original reference or real views).

In general, view synthesis methods can be divided into three categories based on how much geometric information has been used, i.e., whether the method uses explicit geometry, implicit geometry, or no geometry at all [2]. However, a common problem introduced in traditional synthesis methods is that there are some regions in the synthesized view are not visible in the captured views. It is also not easy to obtain the value of the pixel located in these regions, since the information to interpolate the pixel is not available. This problem may cause holes in synthesized views, and also referred to as dis-occlusion problem.

There are several view synthesis algorithms have been proposed to reduce the distortion caused by hole or dis-occlusion problems to improve the quality of synthesized views [3]-[6]. In [3], Yang *et al.* introduced a scheme of reliability reasoning on 3D warping which assesses the reliability of each pixel value in the synthesized view then withdraws the unreliable pixels from the virtual view. Lee *et al.* [4] proposed a background contour region replacement method to clean background noises in the warped views.

In this paper, we proposed an efficient view synthesis method to improve the quality of synthesized views. Unlike traditional methods, the proposed method employs both temporal and disparity correlations between views to reduce the distortion caused by hole or dis-occlusion problems, so as to enhance the quality of the synthesized view image.

The rest of this paper is organized as follows. In Section II, the proposed method is described in detail. Section III shows performance evaluation of the proposed method, and finally, conclusions and future works are presented in Section IV.

II. Proposed View Synthesis Method

Fig. 1 shows a general 3D-TV view synthesis framework of the proposed method. As shown Fig. 1, at the receiver, the 3D-TV decoder extracts and decodes the original real views, left and right, from the received bitstream. Based on these decoded real views, DIBR view synthesis method is utilized for generating intermediate views. Then, by employing a motion compensation based frame interpolation (MCFI), we can obtain temporal prediction view. Finally, the intermediate view are combined with the temporal prediction view using the virtual view fusion to obtain final synthesized view. More details on the performance of the proposed 3D-TV view synthesis are presented in the following subsections.

A. DIBR based view synthesis

As mentioned before, DIBR technique is essential for most 3D-TV applications. This technique allow us to generate intermediate views (or virtual views) based on the received real view sequences and the knowledge of depth information