Texture plus Depth Video Coding Using Camera Global Motion Information

Fei CHENG[1], Tammam TILLO[2], Jimin XIAO[1], Byeungwoo Jeon[3]

1. Xi'an Jiaotong-Liverpool University, Suzhou, China 2. Free University of Bozen-Bolzano, Bolzano, Italy 3. Sungkyunkwan University, Seoul, South Korea

Published on IEEE Transactions on Multimedia

ABSTRACT

In this paper, a novel motion-information-based 3D video coding method is proposed for texture plus depth video, which addresses the issue that the **traditional motion estimation methods cannot work well for all types of motion**.

The global motion information is exploited to :

1. project the reference frames to the view of current-to-be-encoded frame,
2. The motion compensated virtual frame becomes more similar to the current-to-be-encoded frame.

With the computer graphic sequences, for H.264, the average gain of texture and depth coding are up to 2 dB and 1 dB, respectively. For HEVC and HD resolution sequences, the gain of texture coding reaches 0.4 dB.

INTRODUCTION

In many video capturing scenarios, the camera is not static and the image content changes with the camera's global motion. The impact of the camera motion on texture and depth images can be pictorially presented as in the following Fig.1. In this example, a green cube and an orange cylinder are captured by a camera:

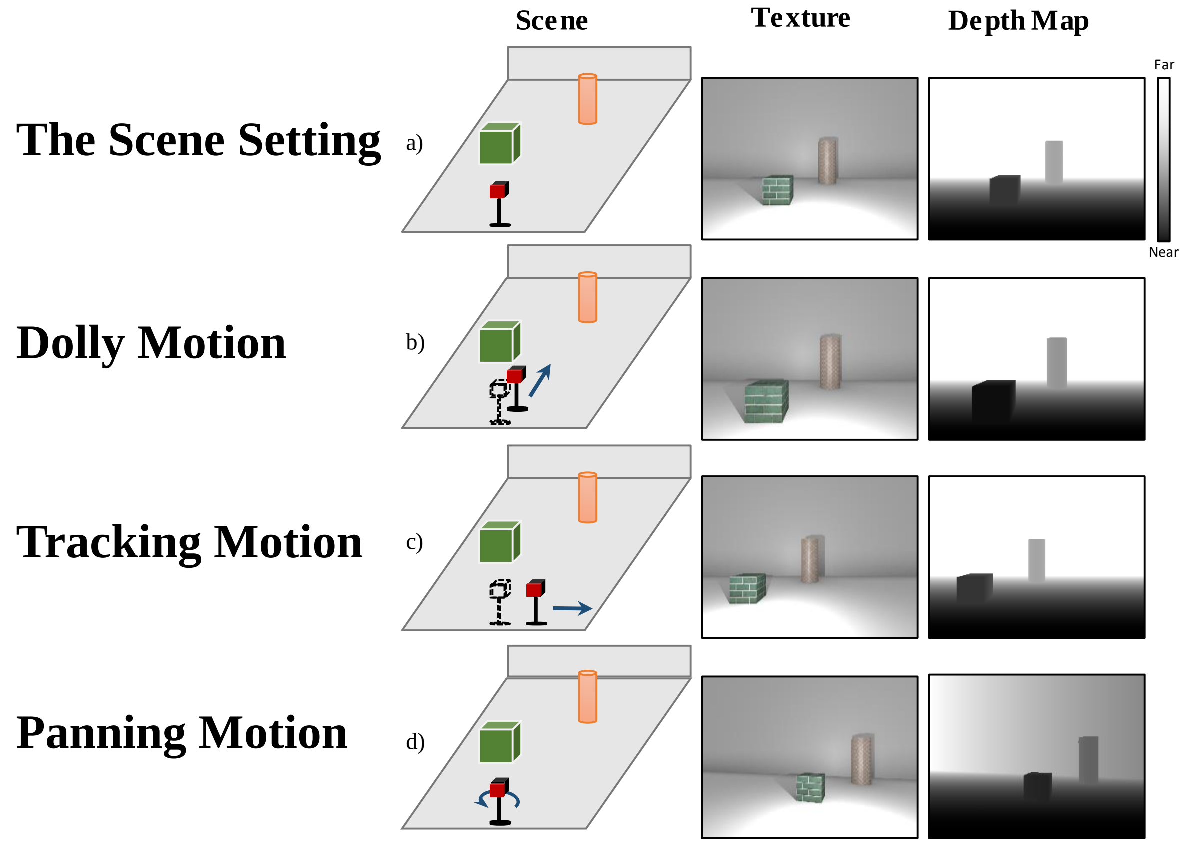

Fig.1 Impact of the camera global motion

The above principle shows that the impact from the global motion is able to be estimated using the camera motion information, which is the key concept of the proposed method.

METHOD

The **synchronously sampled global motion information** of a camera **and its depth map** is exploited to improve the coding performance of both texture and depth sequences.

The block diagram of the proposed motion information based texture plus depth video coding scheme is shown in Fig.2.

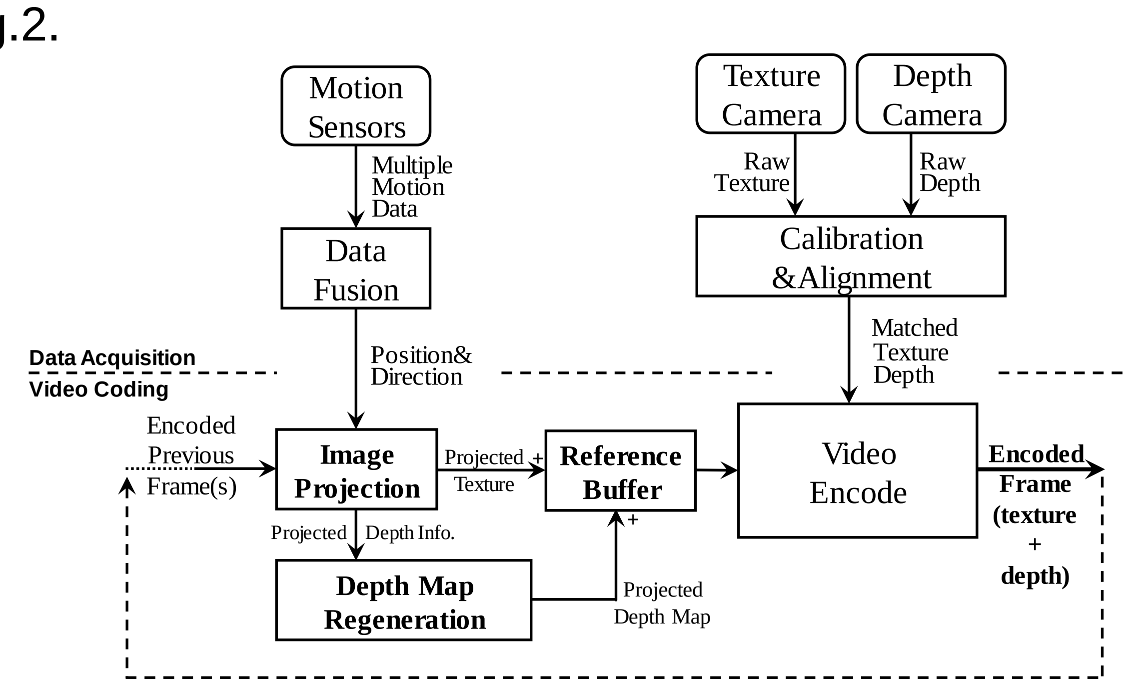

Fig.2 The block diagram of the proposed method

The data acquisition and processing are used to obtain and process the proper video sequence and motion data.

The image projection technique and depth re-quantization method are used in the preprocessing stage to reduce the redundancy between temporal neighboring frames.

The video coding part encodes the texture and depth sequences exploiting the motion information.

In the encoder, the reference buffer mechanism is improved in the method. Fig.3 shows one example that the V'_{n-m} are virtual reference frames projected from previous (or future) frames.

This reference frame buffer mechanism is flexible. As the virtual reference frames cost more memory size, the length of buffer list should be decided based on the performance of the device.

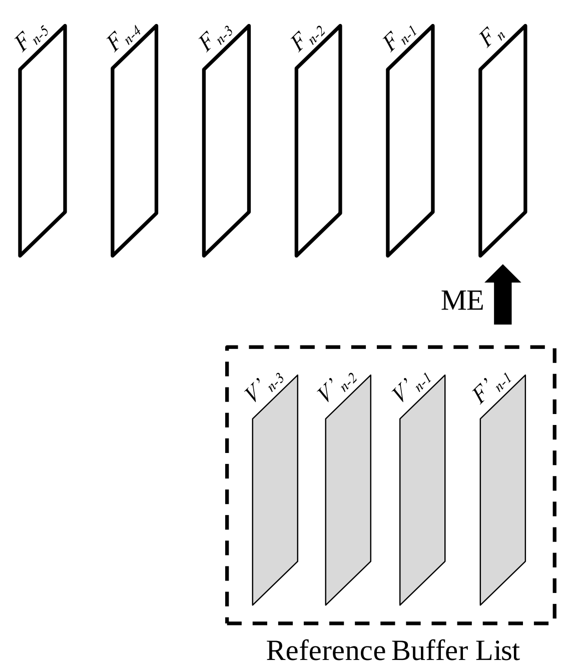

Fig.3 Encoder reference buffer mechanism

EXPERIMENTS AND RESULTS

As there are no proper dataset, we produced texture and depth video sequences including CG and real scenes. Fig. 3 shows an example of a CG sequence.

We tested all the types of basic motions; different resolutions, including VGA and HD; different scenes, such as real and CG scenes; and different codec, including H.264/AVC and H.265/HEVC.

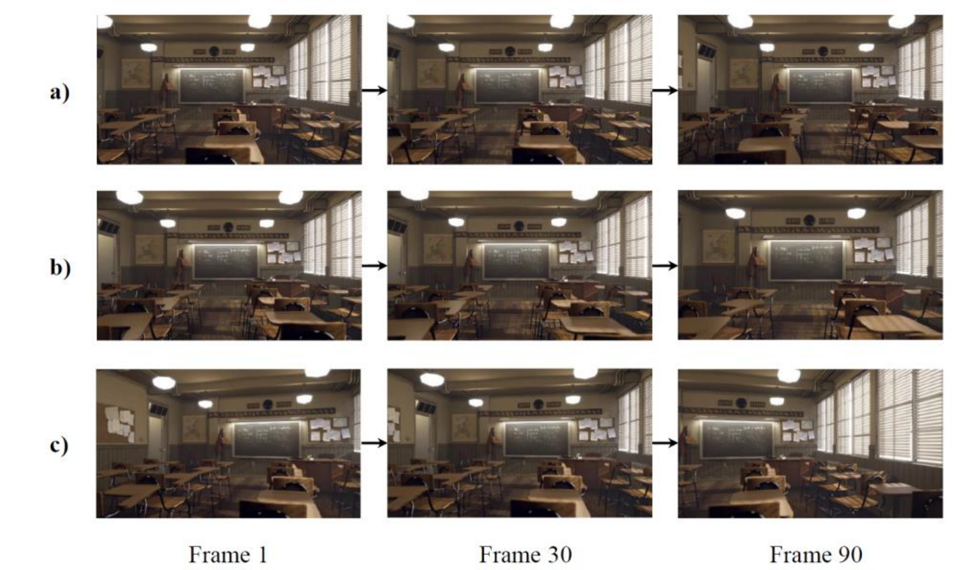

Fig.4 Different motion types of detailed CG scenes used for HD resolution

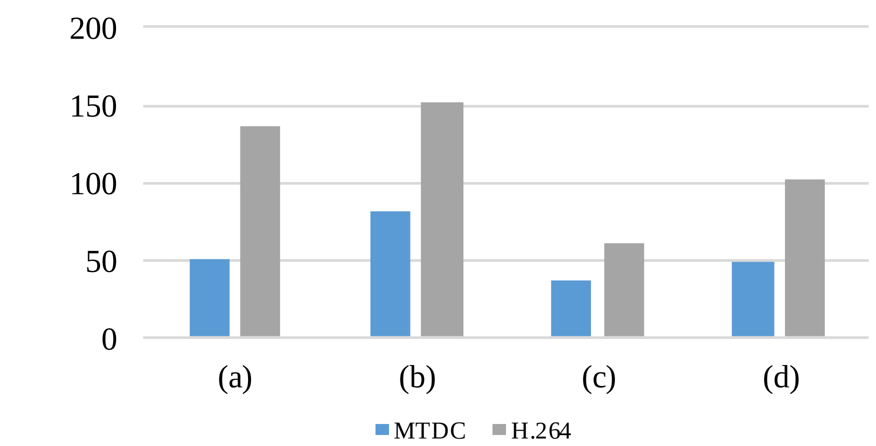

Fig.5 Average number of intra-blocks per P-frame of texture sequences for different speeds of motion using the MTDC scheme and standard H.264/AVC

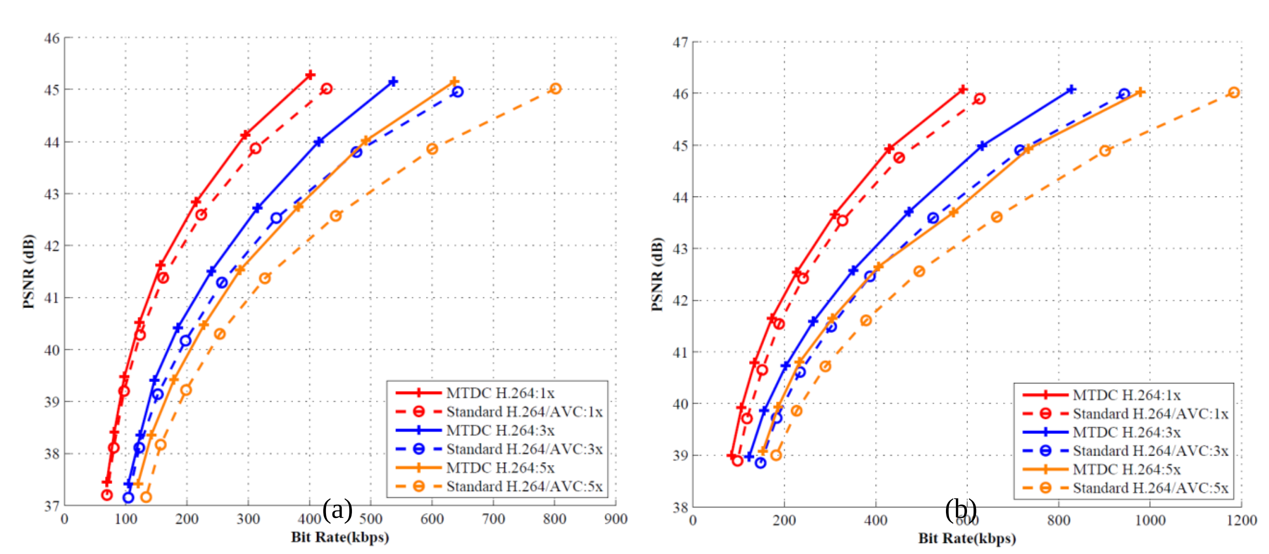

Fig.6 H.264 based results: (a) panning motion (b) dolly motion

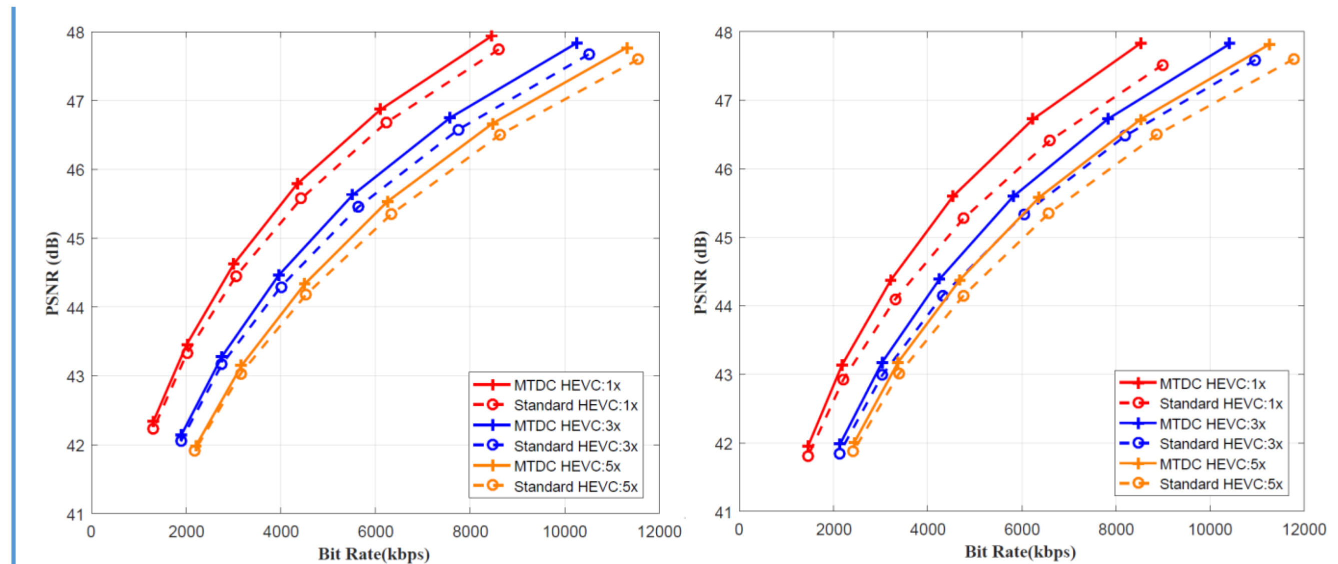

Fig.7 H.265/HEVC based results: (a) panning motion (b) dolly motion

SUMMARY OF RESULTS

The proposed method improved the performance for all the cases

In summary for the H.264/AVC :

1. the VGA resolution CG texture and depth sequence coding performance are enhanced by 2 dB and 1 dB, respectively

2. for the realistic texture video, the performance improvement is around 0.5 dB for texture video and 0.7 dB for depth video.

For H.265/HEVC the MTDC scheme enhances the coding performance :

1. using VGA resolution by up to 0.3 dB

2. using HD resolution sequences by up to 0.4 dB

It is noted that the accuracy of the depth and motion information affects the performance of realistic video coding. With the further quality improvements in depth cameras and motion sensors, the performance of realistic video will be improved and approach the performance of CG video.

ACKNOWLEDGEMENT

This work was supported by National Natural Science Foundation of China (No.61210006, No.60972085, No.61501379), Jiangsu Science and Technology Programme under BK20150375 and in part by the MSIP, Korea, under the GITRC support program (IITP-2017-2015-0-00742) supervised by the IITP.

mmtlab.com