INTERNATIONAL ORGANISATION FOR STANDARDISATION
ORGANISATION INTERNATIONALE DE NORMALISATION
ISO/IEC JTC1/SC29/WG11
CODING OF MOVING PICTURES AND AUDIO

ISO/IEC JTC1/SC29/WG11
MPEG2015/M37232
October 2015, Geneva, Switzerland

Source **Poznań University of Technology**
Status **Proposal**
Title **Depth based view blending in View Synthesis Reference Software (VSRS)**
Author **Krzysztof Wegner, Olgierd Stankiewicz, Marek Domański**

Introduction

At the 112th meeting in Warsaw in July 2015 a Call for Evidence on Free-viewpoint Television (FTV) has been issued [1]. During the works on preparation of CfE on FTV in the third phase of FTV, a software called Enhanced View Synthesis Reference Software [2] has been developed. This software has been thoughtfully tested, both for dense linear arrangements of cameras, and arc camera arrangements. Recent wide angle angular multiview sequences like Poznan Block [3], or Big Buck Bunny [4] revealed some problem with quality of view synthesis. Analysis of current view synthesis algorithm showed that observed problem is caused by view blending stage. This paper proposes an improved view blending mode for enhanced quality of virtual views.

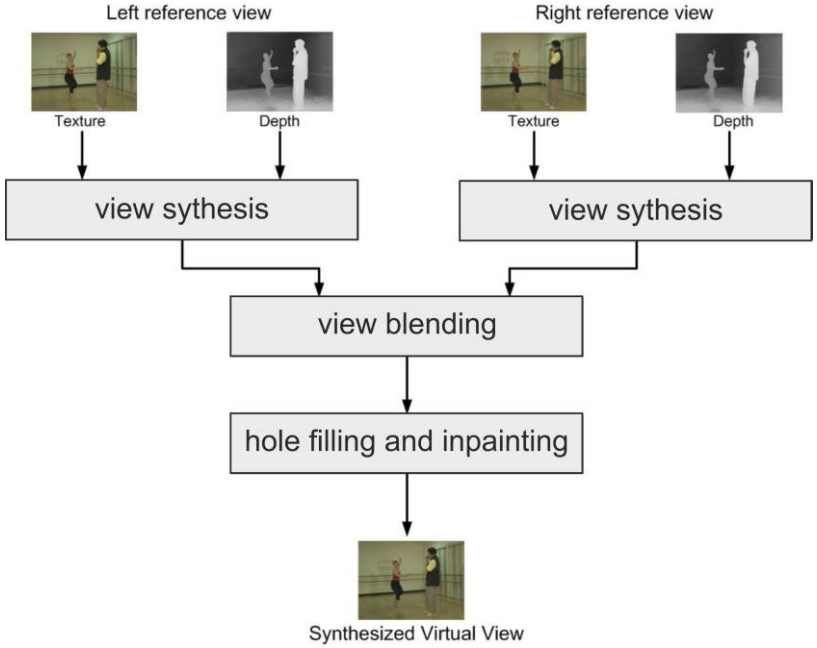

Figure 1. Block diagram of the view synthesis algorithm implemented in VSRS

1. View Blending

1.1. Current approach

In the current version 4.0 of VSRS [2] virtual views are being synthesized based on two input views called reference views (figure 1) . Each input views is used to synthesize independently one version of requested virtual view. As a result, two versions of virtual views called left and right are created. After synthesizing both version of virtual views, both of them are merged together in order to fill occlusions and create single image of the requested virtual view.

Currently in VSRS there are two ways of merging together two intermediate version of virtual views. In the first one, the left and the right virtual views are blended together in proportion α and β related to the distance between given virtual view position and position of two input views used (figure 2). Second method assumes no blending, and just picking one version of virtual view and only filling disoccluded, missed regions with appropriated regions from the second version of the virtual view.

In both modes, the merging operation is performed only based on synthesized images without using depth data

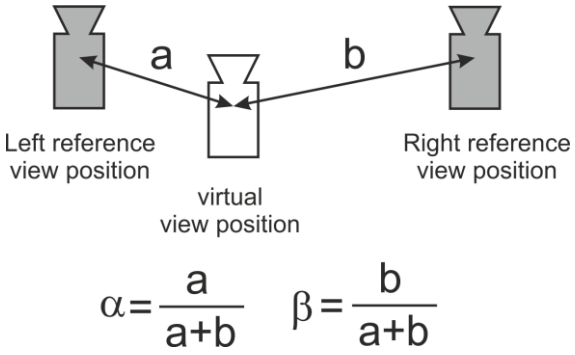

Figure 2. Weights used in view blending in VSRS

1.2. Problem

To illustrate the problem let's assume scene composed of two objects A and B , observed by two cameras as presented in figure 3. Those views are further used for view synthesis as left and right reference views of some virtual view also indicated in figure 1. In depicted situation both objects A and B are visible from the left view, but from the right view only object A is visible. Object B is out of field of view of the right view. If real camera would observe the scene from position of virtual view, both objects would be visible.

Let's consider what happens during view synthesis. During view synthesis process two versions of virtual view (maked in red in figure 3) are synthesized, one based on the left view and one based on the right view. Because of field of view of the left view, virtual view synthesized based on the left view contains both object A and B. But virtual view synthesized based on the right view contains only object A as the object B is not visible from right viewpoint. In virtual view created based on right view, at the position of object B, some background objects are visible (instead of invisible object B which is out of field of view of the right camera)

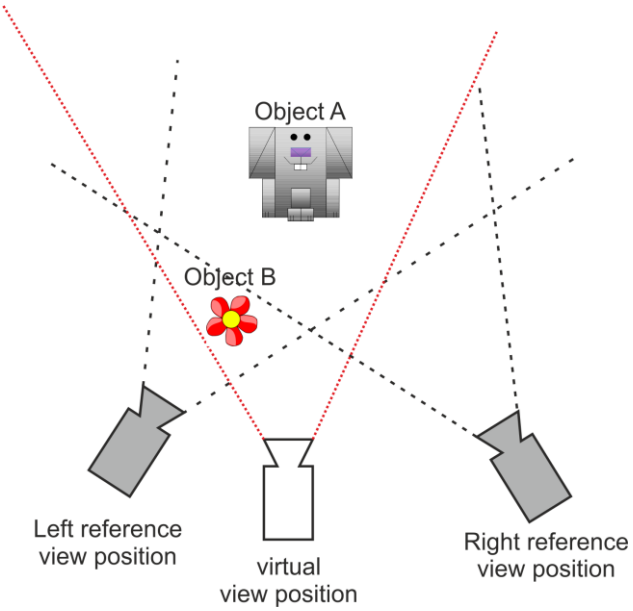

Figure 3. Exemplary scene demonstrating a problem of different fields of views of two reference views used to synthesize requested virtual view.

During final virtual view composition in either of modes, that are currently present in VSRS, there is a problem with area occupied by object B in virtual views. In one version of the virtual view, object B is present and in the other is not. Currently VSRS either tries to blend both views, or arbitrarily selects one of them. In either way resultant virtual views is not correct. In one option object B has wrong colors (as a results of blending of object B with background) or is not present at all (figure 4).


Figure 4. Example of wrongly synthesized views: synthesized virtual view (on the left) and the real one (on the right).

1.3. Proposed view blending method

Proposed solution to above mentioned problem is to analyze the depth during a view blending and selects pixels closer to camera position of the virtual view. Then, during the final composition of a virtual view ,a closer object will occlude further one, and even if in one of the

reference input views some object was not visible, finally, the correct image (pixel) will be selected as a output of the view blending process.

Depth map of a given version of virtual view is being created as a side product of view synthesis in VSRS. For each pixel of a virtual view, the proposed method analyzes depth maps of both version of virtual view and as a final color of given pixel is selected a color from the version of virtual view where pixel was located closer to the requested virtual camera position. Color of a pixel with smaller depth value is selected. If depth values for a considered pixel from both versions of virtual views are similar then final color of considered pixel is generated as a weighted sum of pixel colors from both versions of the virtual view. Weights are calculated exactly in the same way as in view blending mode illustrated in figure 2.

Figure 5. Comparison of view synthesis for proposed depth based view blending (left) with currently used view blending method (right).

Comparison of virtual views synthesized with use of new depth based view blending with current view blending mode is presented in figure 5. Thanks to analysis of depth maps during view blending, flowers on the left part of the virtual view are synthesized correctly.

2. Conclusion

We presented an new depth based view blending mode that solves the problem of different field of view of views used to synthesize virtual views. We recommend to use it as a reference in further development and quality assessment of the new proposals of FTV formats.

Acknowledgement

This work was supported by the funds of National Science Centre, Poland, according to the decision DEC-2012/05/N/ST6/03378.

3. References

[1] [CfE] “Call for Evidence on Free-Viewpoint Television: Super-Multiview and Free Navigation”, ISO/IEC JTC1/SC29/WG11 MPEG2015/N15348 June 2015, Warsaw, Poland.

[2] [VSRS] K. Wegner, O. Stankiewicz, M. Tanimoto, M. Domanski, „Enhanced View Synthesis Reference Software (VSRS) for Free-viewpoint Television”, ISO/IEC JTC1/SC29/WG11 MPEG2013/M31520 October 2013, Geneva, Switzerland.

[3] M. Domański, A. Dziembowski, M. Kurc, A. Łuczak, D. Mieloch, J. Siast, O. Stankiewicz, K. Wegner, „Poznan University of Technology test multiview video sequences acquired with circular camera arrangement – “Poznan Team” and “Poznan Blocks” sequences”, ISO/IEC JTC1/SC29/WG11 MPEG2015/M35846 February 2015, Geneva, Switzerland.

[4] P. T. Kovacs, A. Fekete, K. Lackner, V. K. Adhikarla, A. Zare, T. Balogh, “Big Buck Bunny light-field test sequences”, ISO/IEC JTC1/SC29/WG11 MPEG2015/M36500 June 2015, Warsaw, Poland.