

PUCIT

Punjab University College of Information

Technology

**Final Project Submission**

**Version 1.0**

**Depth Image Based Rendering Library**

Syed Saqib Rafi Hasanie, Umaima Zulfiqar, Arslan Ahmad

*College of Information Technology, University of the Punjab*

## Abstract— 3D video has gained a lot of attention over the past decade due to its vast applications in film, television, animation and virtual reality. The design of intermediate view synthesis algorithms that are efficient in terms of both computational complexity and visual quality is a primary goal in the fields of 3D free-view TVs and displays. The Depth Image-Based Rendering (DIBR) approach is introduced with three different advanced methods. The DIBR algorithm can be used in 3-D video applications to synthesize several different perspectives of the same scene, for example, from a multiview-video-plus-depth (MVD) rendering. This MVD format consists of video and depth footage for a limited number of original camera views of the same natural scene. Here, DIBR methods allow the calculation of new additional views. An inherent problem with the view synthesis concept is the fact that the image information that is occluded in the original views can be made visible, especially in views extrapolated beyond the range of vision of the original cameras. The presented algorithms synthesize these occluded textures. Detailed experiments show significant objective and subjective gains from the proposed method compared to more modern methods.

## 1. INTRODUCTION

## Depth perception in human visual system is a complex process. Due to interpupillary distance (IPD, left and right eyes receive two slightly different images of the 3D scene which results in positional differences of the objects. These positional differences (usually in horizontal direction), referred to as binocular disparities, are processed in the visual cortex of the brain to yield depth perception. This phenomenon can be simulated artificially by presenting the eyes two images of a 3D scene capture at slightly different viewpoints and it is called stereoscopy. Today, many electronic companies are manufacturing 3D-TVs and are aiming to produce high quality 3D televisions at affordable prices.

## “Autostereoscopy” is the 3D display technology that does not require the viewer to wear any special eyeglasses to visualize the 3D display. It is very sensitive to the position of the viewer. To enjoy the 3D experience the viewer must be in a predefined area in front of the display. The autostereoscopy lacks the horizontal parallax that is when the viewer moves horizontally while viewing the display, he will see the same 3D scene which is quite unnatural.

## “Multiview autostereoscopy” provides horizontal parallax by rendering more than 2 views, up to 50 views or more. As a result when the viewer roams around the display he experiences a new set of views in each viewing zone giving him a natural 3D experience. Although no stereoscopy based 3D display provides vertical parallax. That is, the viewer cannot move up or down to see the 3D scene from different vertical angles. Moreover, the multiview autostereoscopic displays provide a limited horizontal parallax around 20-30 degrees, which is still far from real 3D experience.

## There are many other technologies introduced like Integral Imaging, Holography and Volumetric Displays, those are relatively costly and have backward compatibility issues although they still being use but the stereoscopy and multiview autostereoscopy have established as the favorite 3D technologies due to low cost , ease of data acquisition and backward compatibility.

## To enable multiview autostereoscopy and to provide a reasonable horizontal parallax, a huge number of views must be presented at the receiver side for this a technique called Depth Image Based Rendering is introduced. However, capturing, coding and transmitting such a large number of views are not practical due to various cost, hardware, and bandwidth constraints. Therefore, few camera views are captured and transmitted whereas a large number of views are synthesized in real time on the receiver side.

## Depth Image Based Rendering is the view synthesis technique that exploits the depth maps to warp the available views to virtual viewpoints to generate a novel view of the 3D scene. The below techniques are used to create a library for placing different Depth Image Based Rendering algorithms for user’s convenience. This library contains the number of efficient codes. The codes are evaluated on the basis of efficiency, and after a proper analysis of algorithms.

## 2. DEPTH IMAGE BASED RENDERING WITH INVERSE MAPPING

## Abstract— This article focuses on the design of a low complexity view synthesis algorithm that produces better quality of the virtual image. A new view synthesis technique is proposed to create a virtual view from two video sequences with the corresponding depths. The technique employs low complexity integer pixel precision deformation and a novel approach to hole filling based on reverse mapping. The proposed technique [1] is tested in various video sequences and compared with existing state-of-the-art methods, obtaining excellent results both in terms of signal-to-noise ratio and visual quality.

## 2.1- INTRODUCTION

## All DIBR techniques comprise two main algorithmic steps; first, the intermediate virtual view is generated by warping the original pixels to the corresponding intermediate positions, based on depths and camera parameters. After that, two warped views are blended to create a single virtual view. Because of geometrical occlusions, depth imprecision and/or lossy coding of depth there may still be some pixels of the virtual view which are uninitialized, referred to as holes. In DIBR, the second step consists in estimating such missing pixels usually exploiting image inpainting techniques.

## This major contribution of this paper is a novel low complexity View Synthesis with Inverse Mapping (VSIM). VSIM achieves excellent results applying warping with integer pixel precision and recovering synthesis holes using a novel inverse-mapping function.

## 2.2- PROPOSED VIEW SYNTHESIS WITH INVERSE MAPPING (VSIM) TECHNIQUE

## The proposed VSIM algorithm works in two steps like other view synthesis algorithm. It takes two views with depths (left view and right view) as input and computes the intermediate virtual view. Each input view is warped to the intermediate position and the two resultant virtual views are merged together to get a single virtual view. While warping the original views to intermediate positions, a mapping function is defined against each warped view. The missing pixels are then computed with the help of the mapping functions. The proposed algorithm assumes the usual horizontal shift camera setup.The first step in inverse mapping is to determine the original locations of holes in holesV by trying to interpolate the missing values in the table shiftL (shiftR). To this end, VSIM applies a median filter to shiftL and shiftR. Not all the pixels of the virtual view can be recovered for three main reasons that are recalled in the following. Holes due to

## Inaccurate depth

## Cracks

## Disocclusion

## The locations recovered by applying median filter to shiftL and shiftR are used to determine the pixels of the original views that can be copied to fill holes of the virtual view. Due to the limited size of the median filter some holes cannot be recovered with this mechanism. We observed that iteratively increasing the size of the filter to recover all holes generally yields poor shift interpolation results with the creation of artifacts in the virtual view. We have found that the few remaining holes can be recovered by simply assuming that their depth is the same as the co-located pixels in the original views.

## 2.3- EXPERIMENTAL EVALUATION

## The proposed VSIM technique has been tested over a number of standard test sequences. VSIM implementation takes as input two views and their depth maps in YUV(4:2:0) format and estimate the intermediate view in the same format. The warping phase works using integer pixel precision, i.e. by rounding to nearest integer all shifted column positions. The chroma components U, V are warped in their native (downsampled) resolutions as well. The proposed inverse mapping procedure is then used to fill the holes, independently on each of the 3 components.

## 2.4- CONCLUSIONS

## In this paper we have presented VSIM, a novel depth image based rendering algorithm that conjugates limited computational complexity and high quality view synthesis results. These conflicting goals have been achieved by using only integer pixel warping on the one hand, and improving the visual quality of the hole filling algorithm on the other hand. Hole filling is based on the novel idea of inverse mapping that consists in retrieving the missing pixels from the original views rather than interpolating them from the surrounding neighborhood.

## 3. Depth Image-Based Rendering With Advanced Texture Synthesis

## **Abstract**— In our modern times 2D views of movies, games, televisions and technologies have been revolutionized to 3D views. Now 3D videos are encouraged over 2D videos because of immense user demand. With the increase in demand different techniques for 2D to 3D conversion have been evolved but the problem is all those techniques are not available together. This method [2] is about gathering different Depth Image Based Rendering (DIBR) Techniques at one place which can synthesize a number of different perspectives of the same scene, to facilitate the users in order to create 3D videos easily.

## 3.1- INTRODUCTION

Texture synthesis is an appropriate technique to fill missing image regions with known information. This method works in parametric or non-parametric modes. Parametric synthesis approaches generate new textures using a compact model with a fixed set of parameters or dynamic. Nonparametric approaches, on the other hand, typically formulate the texture synthesis problem based on Markov Random Field Theory (MRF). Nonparametric approaches can be classified into patch or sample based methods. Sample-based algorithms update synthetic texture as a sample, while patch-based approaches apply a patch update that is, a set is updated of samples simultaneously. Nonparametric synthesis approaches generally produce better paint results than parametric algorithms, and can also be successfully applied to a much larger variety of textures. For the restoration of small and fairly homogeneous regions, painting approaches that rely on solving partial differential equations (PDE) are often used.

In this method, a new approach to managing disocclusions in Synthetic views are presented for 3-D video. The method is based on non-parametric texture synthesis. Statistical dependencies between different images of a sequence are taken into account through a background sprite (BG). Robust initialization

gives an estimate of the unknown image regions that is refined during the synthesis stage.

**3.2- PROPOSED VIEW SYNTHESIS FRAMEWORK**

The proposed framework for generating virtual view with consistent texture synthesis over time is discussed. The texture images and associated depth maps (DM) of an MVD sequence are taken as input. The DMs receive the test data. Next, a projection is made from the original views to the virtual views based on the where the foreground (FG) and background (BG) objects are deformed separately. First, the disocclusions in the DM are filled. Then the BG sprite is updated with the original BG data and the holes in the current image are updated from the BG sprite. Then the remaining holes are initialized and refined with texture synthesis. Finally, the BG sprite is updated with the new synthesized texture.

**3.4- CONCLUSIONS**

The performance of the proposed view synthesis algorithm is assessed with PSNR and SSIM. For the presented results, PSNR is computed locally, that is, only for the defective area in the image, while SSIM is determined for the entire image as it cannot be easily applied to arbitrary shaped regions. SSIM is

provided in addition to PSNR, since the use of PSNR is difficult in case of geometric distortions.

## 4. EFFICIENT DEPTH IMAGE BASED RENDERING WITH EDGE DEPENDENT DEPTH FILTER AND INTERPOLATION

## Abstract- An efficient representation based on depth images with edge dependent depth filter and interpolation is proposed. The proposed method [3] can solve the hole filling problem in the DIBR system efficiently with high quality. The PSNR of the proposed method is better than the previous work. And the subjective view shows that the quality is better. In addition to that, the number of instructional cycles is 3.7 percent compared to the previous job.

## 4.1- INTRODUCTION

In this article, we propose an efficient representation based on depth images with edge-dependent depth filter and interpolation. The location and size of the hole are detected before deformation of the 3D image. And the edge dependent Gaussian filter is used to reduce the hole size efficiently. We then use edge-dependent interpolation to fill the small holes, so the edge information from the synthesized video is preserved.

## 4.2. PROPOSED METHOD

The main problem we want to solve in the DIBR system is to reduce the number of large holes while maintaining the PSNR and subjective quality of vision. Also, the calculation time is too long in the previous methods. Therefore, the proposed method also targets the computationally conscious DIBR system.

There are three characteristics in the proposed method: edge-dependent depth filter, edge-oriented interpolation, and vertical edge rectification.

## 4.3. EDGE DEPENDENT DEPTH FILTER

The proposed method detects the location of the hole using a special edge filter. And the depth map preprocessing is operated on the detected edges. For the left view, we smoothed the low-to-high sharp depth transition in the horizontal direction because only the low-to-high transition would result in holes. For a correct view, we smoothed the transition from high to low acute depth in the horizontal direction because only this transition would result in holes. Using this method, we can reduce the appearance of large holes and have a good quality in the subjective view and PSNR in the horizontal direction because only this transition would lead to holes. Using this method, we can reduce the appearance of large holes and have a good quality in subjective view and PSNR.

**4.4- Edge Oriented Interpolation**

The edge dependent interpolation method is proposed to preserve the edge information of the interpolated area. The average filter is commonly used to fill holes, but cannot preserve the edge information of the interpolated area. We interpolate a small hole along the edge. The information of the interpolated area is preserved. The edge dependent interpolation method detects the minimum intensity difference of the four directions and interpolates the center hole with the average pixel value of the two pairs with a minimum difference.

**3.3. Vertical Edge Rectification**

Vertical edge grinding will provide depth precision along the vertical line. If a vertical line with an inconsistent depth value is provided along the line, the line in the warped view will branch.

**4- Data Set and Quality Measures**

For evaluating the proposed algorithm, four MVD test sequences are used: “Book arrival” (S1, 100 frames), “Lovebird1” (S2, 150 frames), “Newspaper” (S3, 200 frames), and “Mobile” (S4, 200 frames). S1, S2, and S3 have a resolution of 1024 \* 768 samples, while S4 has a resolution of 720 540 samples. For each sequence, the rectified videos of several views with slightly different camera perspectives are available. The baseline between two adjacent cameras is approximately 65 mm for all test sequences. We consider one or two original—but not necessarily adjacent—cameras (left and right view) to assess the performance of our approach.

**4. RESULTS**

Results have been simulated by software with C language. Subjective view, CPU instruction cycles, misdetection number, hole number reduction rate

**5- Conclusion**

We have discussed three different approaches to do hole-filling. PSNR values of DIBR with Inverse mapping is attached.

**6. REFERENCES**

[1] Muhammad Shahid Farid, Maurizio Lucenteforte,MarcoGrangett “Depth Image Based Rendering with InverseMapping”

## [2] Patrick Ndjiki-Nya, Member, IEEE, Martin Köppel, Dimitar Doshkov, Haricharan Lakshman, Philipp Merkle, Student Member, IEEE, Karsten Müller, Senior Member, IEEE, and Thomas Wiegand, Fellow, IEEE, “] Depth Image-Based Rendering With Advanced Texture Synthesis for 3-D Video”

## [3] Wan-Yu Chen, Yu-Lin Chang, Shyh-Feng Lin, Li-Fu Ding, and Liang-Gee Chen, “EFFICIENT DEPTH IMAGE BASED RENDERING WITH EDGE DEPENDENT DEPTH FILTER AND INTERPOLATION”

DSP/IC Design Lab