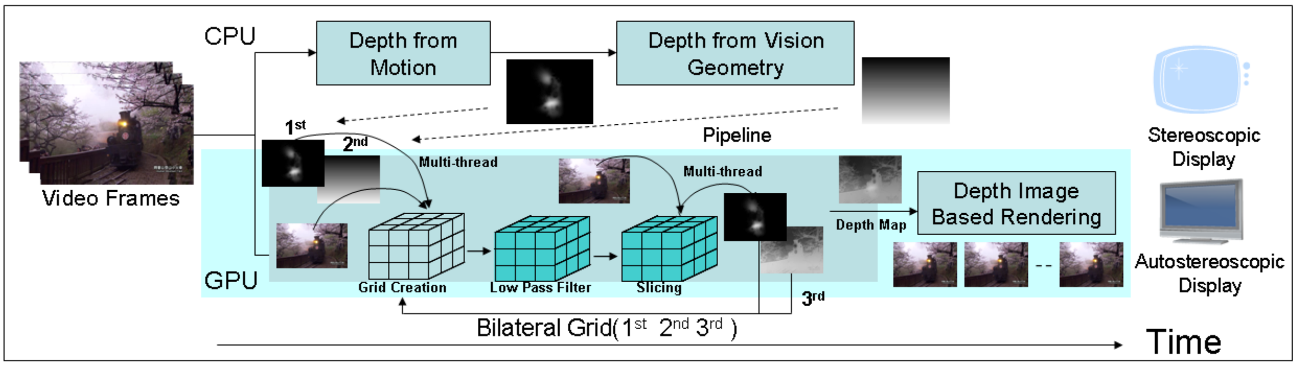
A Real-time Video 2D-to-3D with the Bilateral Grid

Chun-Te Wu, Wei-Hao Huang , Chih-Hao Liu, Wei-Jia Huang, Kai-Che Liu and Ludovic J. Angot

Electronic and Opto-Electronics Research Laboratories, Industrial Technology Research Institute, Taiwan, ROC.



1. Introduction

The new data structure, the bilateral grid, was presented by Jiawen et al. to make bilateral filter algorithm become simple implementation. Based on the data structure, the GPU CUDA-based optimization is proposed to have more efficiency in using GPU shared memory and massive multithreading. Meanwhile, a commercial application, the video 2d to 3d conversion which was presented by Ludovic et al. is also re-designed by applying the proposed CUDA-based bilateral grid three times to obtain better 3D quality in real-time. Depth map are created and modified by adjusting bilateral grid parameters.

2. Bilateral Grid on CUDA

NVIDIA has presented CUDA architecture to program in C language which can be used to arrange multithreading and control shared memory in GPU. Based on the architecture, the GPU CUDA-based bilateral grid is optimized in three steps, grid creation, low-pass filtering and slicing, respectively.

For grid creation, memory allocation in global memory is replaced by shared memory to achieve better performance. During mapping pixels into grid voxels, massive multithreads are created and each thread processes one pixel. Moreover, multithreads are arranged based on the grid voxel structure which means a GPU block in shared memory contains all threads for the same grid voxel. Data layout is designed as three-dimensional array structure by using 32-bit floating point so that moving data is much faster because of coalescing property. After grid creation, the three-dimensional Gaussian filter is separated into three one-dimensional low-pass filters. A GPU block processes one direction filter in x-y plane so that several blocks can work simultaneously and then convolute the z direction. The third step is slicing, which 2D result is interpolated from grid. Based on the CUDA 3D hardware texture filtering, the tri-linear interpolation and normalization are implemented.

3. Video 2D-to-3D Application

In the previous proposed work in Angot et al., the bilateral filter was used to filter initial depth map so that a 2D video can be converted into 3D content represented by 2D video with depth video. However, the depth map quality was not good enough to describe depth of moving objects. Therefore, the proposed algorithm is shown as Fig. 1 for video 2d-to-3d conversion based on applying bilateral grid three times with different parameters.

First, the motion map is calculated from video frames to select initial depth map from several default geometry descriptions. For example, when there's motion in the surrounding area, the initial vanishing point will be put in the center. The idea is to locate vanishing point based on the motion distribution. After obtaining motion map and initial depth map, bilateral grid is applied to filter these two maps with two different spatial (ss) and range (sr) sampling rate parameters, where ss = 32 and sr = 16 for creating segmented initial depth map, and ss = 8 and sr = 4 for modifying motion map with segmentation, respectively. In these cases, original 2D frame is the reference image. Based on the proposed CUDA-based bilateral grid architecture, filtering two maps only need 9ms. Following the bilateral grid is applied again to modify the segmented initial depth map by using segmented motion map as reference image. Since the purpose is to make edges sharp to enhance moving objects and its segmentation, the segmented motion map is used to filter segmented initial depth map with small parameters, such as ss = 8 and sr = 4. Because of applying bilateral grids three times with different parameters to create and modify the depth map, motion and geometry properties are fusion to represent the 3D information.

In the video 2d-to-3d application, the proposed algorithm is optimized based on CPU-GPU double buffering pipeline architecture. Motion map and initial geometry depth map are created on CPU, meanwhile, GPU is designed for multithreading bilateral grid and depth image based rendering (DIBR) to stereoscopic or autostereoscopic display. Based on the proposed CPU-GPU pipeline, CUDA-based bilateral grid and GPU parallel optimized DIBR architecture, for converting 2D video to 3D and rendering to 1920x1080 9-views autostereoscopic display, 30fps real-time performance is achieved by using Intel Core 2 Duo 3GHz CPU and NVIDIA GeForce GTX 280 GPU. Applications include 3DTV, 3D video conference and 3D Movie.

References

Jiawen Chen , Sylvain Paris , Frédo Durand, Real-time edge-aware image processing with the bilateral grid, ACM Transactions on Graphics (TOG), v.26 n.3, July 2007

Ludovic J. Angot, Wei-Jia Huang, Kai-Che Liu. 2010. A 2D to 3D video and image conversion technique based on a bilateral filter. Proceedings of SPIE Vol. 7526.

Copyright is held by the author / owner(s).
SIGGRAPH 2010, Los Angeles, California, July 25 – 29, 2010.
ISBN 978-1-4503-0210-4/10/0007