A Real-Time 1080p 2D-to-3D Video Conversion System

Sung-Fang Tsai, Chao-Chung Cheng, Chung-Te Li, and Liang-Gee Chen
Graduate Institute of Electronics Engineering, National Taiwan University

***Abstract*—In this paper, a 2D-to-3D video conversion system capable of realtime conversion of 1920×1080p 2D video to 3D video is presented. System fuses global and local depth generation modules to generate depth image, and use depth image based rendering(DIBR) algorithm to render 3D video. The system is implemented both on software and hardware. Software is based on multi-core system with CUDA platform. To optimize performance, several techniques are proposed, including unified streaming dataflow, multi-thread schedule synchronization, and CUDA acceleration. Real-time 1920×1080p 2D-to-3D video conversion system running at 30fps is achieved.**

I. Introduction

3D video are getting immense public attentions recently because of the vivid stereo viewing experience. However, most of the existing content do not have depth information. In these cases, 2D-to-3D conversion on the fly is required for 3D video playback system. From human perceptual and visual knowledge, depth cues can be generated accordingly. Previous proposed system uses different cues for depth generation [1] [2]. Our latest system fuses global and local depth cues from video analysis to generate depth information and renders the generated video to 3D display [3]. Real-time performance for 1920×1080p conversion is achieved in this paper with proposed unified streaming dataflow, multi-thread schedule synchronization, and CUDA acceleration.

II. Proposed System

A. Algorithm

In this paper, we generates depth maps using the human visual perception characteristics of luminance and color. The method has two major parts, edge feature-based global scene depth gradient and texture-based local depth refinement.

In the global depth map generation, we accumulate the edge feature count of the frame horizontally to get the horizontal complexity of the frame. The near-to-far global scene depth is generated by analyzing the edge complexity on each row.

The local pixel value, Y, Cb, and Cr component of the video content is used to refine the detail depth value. Each color component is mapping to a linear gain from 1-Cth to 1+Cth.

Combining the global depth gradient and local depth refinement, the depth map has a comfortable and vivid quality.

B. System Optimization

In proposed 2D-to-3D conversion system, videos for both left and right views are generated from original single view video. The system consists of two major parts: depth image generation and depth image based rendering as shown in Fig. 1. Scene cues and texture cues are used in this implementation. After the depth image is generated, the left and right views are then rendered with depth image based rendering (DIBR) algorithm. The system is implemented as an DirectShow transform filter and integrated with 3D video player for realtime demonstration.

The performance bound is memory bandwidth. Execution path of depth generation is complex and can be run on Multiple Instruction stream, Multiple Data stream(MIMD) architecture. Depth image based rendering is rather fixed and is more suitable on Single Instruction stream, Multiple Threads (SIMT) architecture. As a result, we put depth generation on CPU and apply bandwidth control techniques and DIBR on GPU to optimize performance. The following section will present the proposed method for performance optimization.

1) Unified Streaming Dataflow for Multi-Format Processing: In proposed system, the video decoder may output video in various formats, color space, and chroma subsampling modes. The default color converting in Directshow will cause page thrashing and the CPU will stall on data transmission. To avoid this problem, an unified streaming dataflow is proposed here. Filter parameter for cue generation is projected into the input color space. For various pixel format and packing, we uses specific parameter sets of skip, stride, and span. As a result, all the computation is completely done on the input color space. Whole frame color conversion is saved.

2) Multi-Thread Schedule Synchronization for Data Locality Optimization: Without proper alignment of data locality and thread scheduling, multithreading performance will be limited by bandwidth. Multi-thread schedule synchronization is proposed to adjust processing order and improve data locality. Synchronization points are evenly placed on the input image. At each synchronization point, each thread will check maximum displacement of synchronization point among threads. If it is larger than a threshold value, data loading is postponed to avoid bandwidth overhead. Thread will sleep temporarily to wait for other threads. The bandwidth overhead is reduced with this method.

3) GPU Acceleration for DIBR: Since the algorithm used for depth image based rendering (DIBR) is a highly parallelizable and the processing order is regular, we implement it on GPU using CUDA. Each row is divided into different threads and generate left and right views using depth image. Since the rendered 3D image is on the GPU, it can be directly rendered.

978-1-4244-8712-7/11/$26.00©2011 IEEE

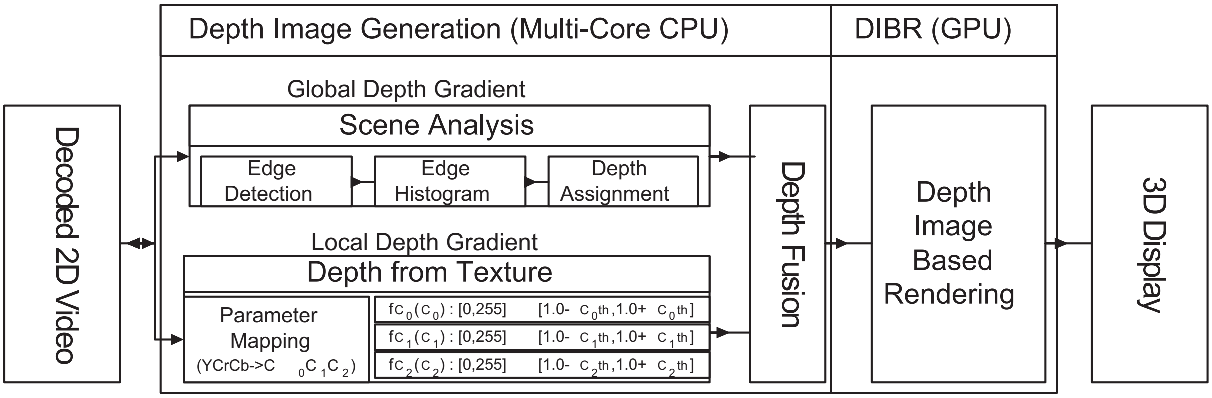

Fig. 1. System architecture for proposed 2D-to-3D video conversion system

Fig. 2. Original 2D images (upper left), depth maps (bottom left), and red-cyan images (right) of test sequences.

III. Implementation Results

The system is implemented on ASUS G51J 3D Notebook with Intel Core i7 CPU 720QM and NVIDIA GeForce GTX 260M. The system is integrated in video player for evaluation. The original video, converted depth map, and converted red-cyan images are shown in Fig. 2. The proposed algorithm running on multi-core CPU can achieve 30fps at 720×480p. With proposed optimization techniques over CPU, 960×540p @ 30fps is achieved. With GPU acceleration for DIBR, 1920×1080p @ 30fps video conversion is achieved since DIBR can be run on GPU efficiently. The real-time performance is verified.

The proposed system is also verified in hardware. The technology used is UMC 90 *nm* 1P9M CMOS. The spec is up to 30fps for 1920×1080p 2D-to-3D conversion. The hardware cost of the proposed system is listed in Table I as a reference.

TABLE I
HARDWARE COST OF THE PROPOSED 1080P 2D-TO-3D CONVERSION SYSTEM

Module	Logic Gates(2-input NAND gate)
Depth Generation	14.5k
DIBR	15k
Control	3k

IV. Conclusion

This paper presents an real-time 1080p 2D-to-3D conversion system. The proposed system combines two major depth generation modules, and optimize system bandwidth and performance by unified multi-format dataflow, multi-thread schedule synchronization, and CUDA acceleration. Hardware version is also implemented as a reference. This work is suitable for 2D-to-3D video playback system.

References

[1] Chao-Chung Cheng, Chung-Te Li, Po-Sen Huang, Tsung-Kai Lin, Yi-Min Tsai, Liang-Gee Chen, "A block-based 2D-to-3D conversion system with bilateral filter," International Conference on Consumer Electronics(ICCE), 2009.

[2] Chao-Chung Cheng, Chung-Te Li, and Liang-Gee Chen, "A 2D-to-3D Conversion System Using Edge Information," IEEE International Conference on Consumer Electronics(ICCE), 2010.

[3] Chao-Chung Cheng, Chung-Te Li, and Liang-Gee Chen, "An Ultra-Low-Cost 2D-to-3D Conversion System," International Symposium, Seminar, and Exhibition of Society For Information Display(SID), Display Week 2010.