Dense Stereo-based Real-time ROI Generation for On-road Obstacle Detection

Soon Kwon[1,2] and Hyuk-Jae Lee[1]
[1]Department of Electrical and Computer Engineering, Seoul National University, Seoul, Korea
[2]Convergence Research Center for Future Automotive Technology, DGIST, Daegu, Korea
{soonyk, hjlee}@capp.snu.ac.kr

***Abstract*—The use of 3D visual information has become widespread as an essential cue for detecting on-road obstacles in ADAS. In this paper, we propose an accurate dense-stereo-based system for the generation of on-road obstacle ROIs. To balance the concerns of computation overhead and algorithm accuracy, this paper presents an efficient depth map generation that combines global stereo matching with depth up-sampling. The entire system has been implemented with a hardware and software partitioning method running on an FPGA and embedded CPU for real-time processing. The implementation results verify that the proposed stereo vision system efficiently outputs accurate ROI candidates for on-road obstacle detection.**

Keywords: stereo matching, depth map genration, region of interest, obstacle detection, advanced driver assistance system

I. INTRODUCTION

In the field of advanced driver assistance systems (ADAS) and autonomous driving applications, 3D computer vision methods are widely used to recognize free space and on-road obstacles [1]. Using depth information, the region of interest (ROI) search can be simplified and is more robust than with conventional 2D search methods [1, 2]. For example, a road surface profile used for free space detection can be easily analyzed by applying the V-disparity method with dense depth map information.

However, the computational overhead in a dense stereo matching process, which finds corresponding positions for left and right image pixels, presents challenging issues to be overcome. A number of studies have shown algorithmic or architectural improvements for speeding up the stereo matching process. Compared with algorithmic approaches [3] that deal with computation efficiencies, architectural approaches [4] are more suitable for real-time computation purposes. By considering hardware-friendly algorithm architectures, they can achieve remarkable reductions in computation time.

We illustrate here a real-time stereo-vision-based ROI generation system for the detection of on-road obstacles. The main contribution is in designing the system to use hardware and software partitioning for efficient implementation as an embedded vision system. The paper is organized as follows. Section II introduces the proposed algorithm. In Section III, the result of the implementation of the proposed hardware and software partitioning is described in detail. Finally, Section IV concludes the paper.

II. OVERVIEW OF THE PROPOSED ALGORITHM

Figure 1 shows the proposed dense-stereo-based ROI generation for the detection of on-road obstacles. In the depth map generation step, the rectified stereo image pair is firstly downscaled to reduce the increase in computation cost, which is proportional to the image size. For the downscaled stereo image, a global stereo matching process based on a 2D Markov random field (MRF) generates a smooth and accurate depth map. Finally, the resolution of the depth map result is reconstructed to an original size by using a depth map up-sampling method [5]. Note that we intend to minimize the quality difference between the original depth map and the up-sampled one.

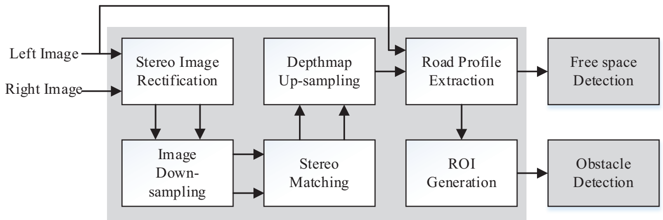

Figure 1. Flow of proposed algorithm

In designing a robust stereo matching algorithm for on-road scenes, we should consider the variation in illumination or the possibility of large textureless regions, which frequently appear in the road or on the vehicle's planar surface. As shown in Figure 2, we use the census transform as a local cost function and iteratively optimize the global matching cost by using a four-level hierarchical belief propagation (HBP) algorithm [3]. Note that the iterative cost update scheme in HBP is effective for the estimation of a good depth value for ambiguous regions such as those mentioned above.

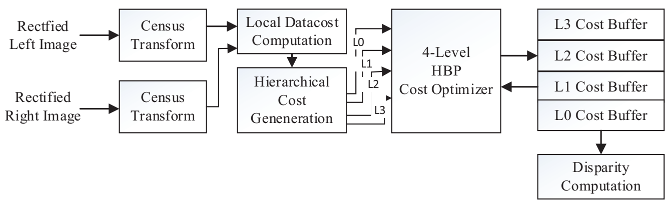

Figure 2. Stereo matching process

In the ROI generation step as shown in Figure 3, the generated depth map and the color image are used to extract the road surface profile and ROI candidates. Using the U-V-disparity method, the initial ROI estimation results are further refined and merged by using scale information and the class-specific features of the obstacles such as the aspect ratio and size.

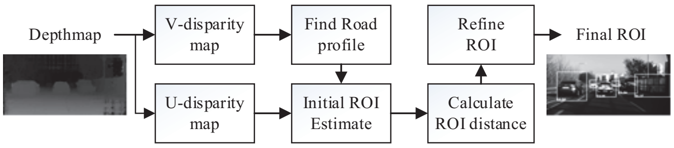

Figure 3. ROI generation

III. IMPLEMENTATION RESULTS AND EVALUATION

In order to overcome the enormous computational overhead of the global stereo matching process, a dedicated hardware accelerator has been designed [4]. Our parallel processing architecture based on hierarchical structures in HBP and using hundreds of processing elements (PEs) is hundreds of times faster than the original HBP algorithm. As shown in Figure 4, we designed an integrated hardware architecture that includes image rectification, cost computation, and HBP-based global optimization logic. The hardware architecture was implemented on the Xilinx Kintex-7 field-programmable gate array (FPGA) chip, and the implementation results are summarized in Table I.

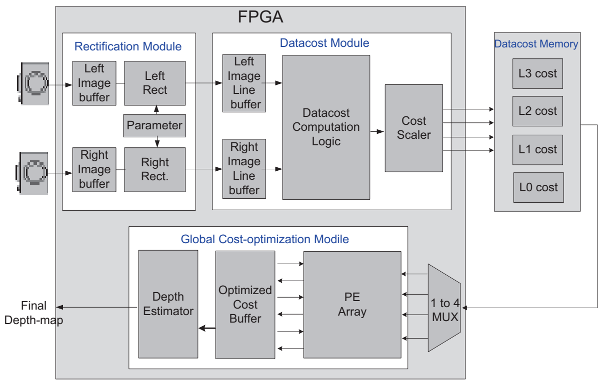

Figure 4. Hardware architecture of stereo matching IP core

TABLE I. RESULTS OF IMPLEMENTATION ON A KINTEX-7

		Implementation Results
clock frequency		53 Mhz
on-chip memory requirement		646KB
block-ram usage	***Total***	68%
	Globalcost	33.4%
	Rectification /localcost	16.3%
	Memory I/F	18.3%
LUT usage	***Total***	62%
	Globalcost	42.1%
	Rectification /localcost	5.7%
	Memory I/F	14.2%

The remaining processes, including image down-sampling, depth up-sampling, and ROI generation, have been implemented in the NVIDIA TX1 embedded processor. Among these processes, the depth scaling processes have been implemented on the TX1 graphics processing unit (GPU) core using the CUDA parallel programming language.

TABLE II. COMPUTATION TIMES FOR REAL-TIME PROCESSING

Process	**Execution Time (ms)**		
	CPU*	***FPGA***	***GPU***
IMAGE DOWN-SAMPLING	3	-	1.5
STEREO MATCHING	2,576	27.0	-
DEPTH UP-SAMPLING	307	-	9.9
Total	**2,886**	**38.4 (26 FPS)**	

* Intel i-7 2.8Ghz CPU

Table II presents the total computation time of the proposed depth map generation system using the FPGA-based hardware accelerator and GPU-based parallel programming. The total time is markedly less than that using only CPU-based execution. Note that the proposed stereo matching process achieves a real-time processing rate of 26 frames per second (FPS) for an HD-sized stereo image sequence.

Finally, the total execution time as executed by hardware and software partitioning achieved a real-time processing rate of 15 FPS. The ROI generation portion consumed only about 5 ms on TX1's CPU core.

IV. CONCLUSION

We have proposed an efficient stereo-vision-based on-road obstacle ROI generation system for use in ADAS. In order to reduce computational costs, an efficient depth map generation process has been implemented. In addition, in order to accelerate the execution, FPGA-based hardware and GPU-based parallel programming are utilized. The implementation results show that the proposed system outputs an accurate dense depth map and multiple ROI candidates in real time.

ACKNOWLEDGMENT

This work was supported by the DGIST R&D Program of the Ministry of Education, Science and Technology of Korea (16-IT-02). It was also supported by the MSIP(Ministry of Science, ICT and Future Planning), Korea, under the ITRC(Information Technology Research Center) support program (IITP-2016-H8601-16-1008) supervised by the IITP(Institute for Information & communications Technology Promotion).

REFERENCES

[1] H. Zhencheng, and K. chimura. "UV-disparity: an efficient algorithm for stereovision based scene analysis." IEEE Intelligent Vehicles Symposium(IVS), 2005.

[2] J. M. Alvarez and M. L. Antonio, "Road detection based on illuminant invariance." IEEE Transactions on Intelligent Transportation Systems 12.1: 184-193, 2011.

[3] P. F. Felzenszwalb and D. P. Huttenlocher, "Efficient belief propagation for early vision", IJCV, 70(1): 41-54, 2006.

[4] S. Kwon, C-H. Lee, Y-C. Lim and J-H. Lee, "A sliced synchronous iteration architecture for real-time global stereo matching." IS&T/SPIE Electronic Imaging. International Society for Optics and Photonics, 2010.

[5] S. B. Lee, S. Kwon and Y. S. Ho, "Discontinuity adaptive depth upsampling for 3D video acquisition", Electronics Letters, 49.25: 1612-1613, 2013.