# A Real-Time Stereo Matching Hardware Architecture Based on the AD-Census

Hyeon-Sik Son<sup>1</sup>, Kyeong-ryeol Bae<sup>1</sup>, Yong-Hwan Lee<sup>2</sup> and Byungin Moon<sup>3,\*</sup>

 <sup>1</sup>School of Electrical Engineering & Computer Science, Kyungpook National University, Daegu, Korea
 <sup>2</sup>School of Electronic Engineering, Kumoh National Institute of Technology, Gumi, Korea
 <sup>3</sup>School of Electronics Engineering, Kyungpook National University, Daegu, Korea
 <sup>1</sup>{soc\_shs1984, puris1}@ee.knu.ac.kr,<sup>2</sup>yhlee@kumoh.ac.kr
 <sup>3</sup>bihmoon@knu.ac.kr

#### Abstract

In this paper, we propose a new stereo matching hardware architecture based on the AD-Census stereo matching algorithm that produces accurate disparity map. The proposed stereo matching hardware architecture is fully pipelined and processes images with disparity level parallelism in real time. Also, it uses modulo memory addressing methods for reducing the size of memory and the usage of hardware resource. The proposed architecture is perfectly synchronized with the input camera clock for real-time performance. Its maximum clock frequency is 197 MHz when it is implemented in an FPGA device.

Keywords: AD-Census, Stereo matching, Disparity map, Real-time

### 1. Introduction

Stereo matching is one of the most actively studied problems in computer vision. Most stereo matching algorithms can be classified into "global" and "local" methods. The global methods minimize energy function with dynamic programming [1], graph cuts [2] or belief propagation [3]. The global methods can successfully suppress the matching ambiguities caused by illumination variation and textureless regions. Thus, these methods generate more accurate results than local methods. However, they require high computational power for optimization process. The local methods compute each pixel's disparity independently of the relationship with neighbor pixels and matching costs are extracted by simple measures such as absolute difference (AD), gradient and Census. The disparity level of each pixel is selected by minimal matching cost. Compared with global methods, the local methods are simpler and have relatively little computation time but they produce less accurate disparity maps.

Also, the matching accuracy and the processing efficiency are major concerns in stereo matching algorithms. Although many algorithms are introduced every year, the accurate stereo methods are usually time consuming [4-5]. For real time stereo matching, some hardware architectures are introduced in [6-7], but these use simple local methods because of the limitation of hardware resources.

<sup>&</sup>lt;sup>\*</sup> Corresponding author.

In this paper, we aim to meet accurate stereo matching requirements with real-time performance by adopting the AD-Census stereo matching algorithm [8] and fully pipelined hardware architecture. The AD-Census algorithm has the combined features of AD and Census transform for more accurate stereo matching.

The rest of the paper is organized as follows. In Section 2, we introduce the AD-Census stereo matching algorithm .In Section 3, we describe the proposed AD-Census hardware architecture. Section 4 presents the experimental environments, and then analyzes the results of the experiments. Finally, we summarize and conclude the paper in Section 5.

### 2. AD-Census Algorithm

The AD-Census is a combined algorithm of AD and Census transform for improved matching accuracy. The Census transform could produce matching ambiguities in image regions with repetitive or similar local structures although it shows the best overall results in local and global stereo matching methods [9]. To overcome this problem, the color information is adapted to alleviate the matching ambiguities.

The AD-Census cost is computed using Equation (1). Given a pixel p=(x, y) in the left image and a disparity level d, the matching cost C(p, d) is computed by combining two simple measures [8]:

$$C(p, d) = min(C_{AD}(p, d), \lambda_{AD}) + min(C_{census}(p, d), \lambda_{census})$$
(1)

where  $\lambda_{census}$  and  $\lambda_{AD}$  are threshold parameters to reject outliers and  $C_{census}(p,d)$  is defined as the Hamming distance for pixel p and its correspondence pd = (x-d, y) in the right image [10].  $C_{AD}$  is the RGB color difference of pixel p and pd:

$$C_{AD}(p, d) = \sum_{i=R,G,B} |I^{i}_{left}(p) - I^{i}_{right}(pd)|$$
(2)

The reason why we employ AD-Census measure is that the AD-Census measure shows improved matching accuracy than individual AD and Census measures. In a recent evaluation by [11], Census produces wrong matches in regions with repetitive local structures, while pixel-based AD cannot handle well large textureless regions. The combined AD-Census measure successfully reduces the errors caused by individual measures.

### 3. Proposed hardware Architecture

Figure 1 shows the block diagram of the proposed AD-Census hardware architecture. It aims at processing depth maps, synchronized with the input camera clock for real-time performance. For this reason, disparity level parallel processing is essentially required. The proposed AD-Census hardware architecture consists of the Image Memory, the Window Selector, the AD cores, the Census cores and the Depth Computation.



Figure 1. Block Diagram of the proposed AD-Census Hardware Architecture

The Image Memory stores the left and right image data from the stereo camera line by line, and it outputs the line image data of the window region. For parallel processing, each line image data need to be stored in each line memory separately. However, it needs only n+1 numbers of line memory devices for left and right image respectively by using modulo memory addressing methods like Figure 2 [12]. Because of modulo memory addressing, rearrangement of the image data is required. The Window Selector aligns the order of vertical line pixels depending on the window region. Each of the AD and Census compute the matching cost for depth computation using its own measure. The Depth Computation computes final matching cost based on the Equation 1 and outputs the depth results by selecting minimum matching cost by tournament comparison.



Figure 2. Modulo Memory Addressing Methods

The AD cores for red, green and blue images have the same structure. Figure 3 shows the AD core for red image processing. The AD core computes the absolute difference of left and right image in disparity level parallelism. For the disparity level parallelism, it has one Left Scan-line Buffer and as many Right Scan-line Buffers as the size of disparity range. The absolute difference values between left and right pixels in vertical line are computed and then the sum of ADs in the vertical line of window region is calculated. Finally, the matching costs of AD are computed by sum of vertical ADs over horizontal line. To reduce the usage of hardware resource, The Scan-line Buffer stores just a vertical line image data of window region. For elimination of bottleneck caused by addition, every addition is computed by using tournament.



Figure 3. Proposed AD Core Hardware Architecture for Red Image



Figure 4. Proposed Census Core Hardware Architecture for Red Image

The Census cores for red, green and blue images have the same structure like the AD cores. Figure 4 shows the Census core for red image processing. The Hamming Weight module makes the hamming weight bit stream through pixel value comparison among the window regions. Also, the Census cores each have as many right hamming weight buffers as the size of the disparity range like the AD cores. The Hamming Distance

module computes hamming distance by bitwise XNOR operation between left and right hamming weights and counts the numbers of '1' and then outputs the hamming distances as Census matching costs.

## 4. Experimental Results

The proposed AD-Census hardware architecture is designed with an HDL and implemented in Xilinx Virtex6 LX760 FPGA. The proposed architecture has 64 disparity ranges and 15x15 window sizes. This architecture is compared with the hardware architecture of [7].

Table 1 shows the FPGA implementation results. Even though the proposed hardware architecture is implemented by combining the AD and Census transform, the hardware usage is similar to that of [7] only with Census transform except for Slice LUTs.

Figure 5 shows the results of AD-Census stereo matching with Middlebury benchmark images [13]. As shown in the Figure 5, the disparity maps of the Middlebury benchmark images are finely extracted. Figure 6 is the disparity map of captured images under common outdoor scenery [14]. As shown in Figure 6, the matching errors still occur due to the occlusion and environmental noise. However, it is possible to recognize obstacles and the distance of them.

The proposed architecture is synchronized with the input camera clock with the maximum clock frequency of 197 MHz.

	FPGA results		Characters
[7]	Slice Registers	53,616	Census transform
	Slice LUTs	60,598	Disparity range 64
	Block Ram/FIFO	192	Rectification included
Proposed	Slice Registers	67,650	AD-Census(AD + Census)
	Slice LUTs	126,850	Disparity range 64
	Block Ram/FIFO	32	

**Table 1. FPGA Implementation Results** 



(a) Teddy image



(b) Corn image



(c) Baby image

Figure 5. Disparity Map Results of Middlebury Benchmark Images



Figure 6. Disparity Map Results of Captured Images under Common Scenery

# 5. Conclusion

This paper proposes a new real-time stereo matching hardware architecture for the high performance and more accurate results. We adopt the AD-Census stereo matching algorithm which can reduce the matching errors caused at individual measures and produce accurate disparity map. The proposed AD-Census architecture is fully pipelined for real-time performance and its operation is synchronized with the input camera clock. It computes accurate disparity of each pixel with disparity level parallelism. According to the experimental results, the proposed stereo matching hardware computes the disparity map in real time. However, the matching errors still occur due to occlusion and environmental noise.

Also, the hardware usage of the proposed AD-Census hardware architecture is similar to that of [7] only with Census transform except for Slice LUTs. The maximum clock frequency of the proposed hardware architecture is 197 MHz in an implementation with an FPGA device, so it is able to support most existing cameras in real time.

In the future, we will study on obtaining more accurate disparity map under outdoor environments and then will implement the proposed AD-Census based stereo matching hardware architecture as an ASIC for the purpose of commercial use. The proposed architecture can be used for various vision applications, such as smart cars, intelligent robots and navigation.

### Acknowledgements

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the ministry of Education, Science and Technology (2011-0013948).

### References

- [1] A. F. Bobick and S. S. Intille, Int. J. Comput. Vision., vol. 33, no. 181, (2003).
- [2] Y. Boykov, O. Veksler and R. Zabih, IEEE Trans. Pattern Anal. Mach. Intell., vol. 23, no. 1222, (2001).
- [3] P. F. Felzenszwalb and D. P. Huttenlocher, Int. J. Comput. Vision., vol. 70, no. 41, (2006).
- [4] A. Klaus, M. Sormann and K. Karner, Editors, "Segment-based Stereo Matching using Belief Propagation and a Self-adapting Dissimilarity Measure", Proceedings of the 18th International Conference on Pattern Recognition, Hong Kong, (2006) August 15-18.
- [5] Q. Yang, L. Wang, R. Yang, H. Stewenius and D. Nister, IEEE Trans. Pattern Anal. Mach. Intell., vol. 31, no. 492, (2009).
- [6] S. Wong, S. Vassiliadis and S. Cotofana, Editors, "A Sum of Absolute Differences Implementation in FPGA Hardware", Proceedings of the 28th Euromicro Conference on Multimedia and Telecommunications Track, Dortmund, Germany, (2002) September, pp. 183-188.
- [7] S. Jin, J. Cho, X. D. Pham, K. M. Lee, S. K. Park, M. Kim and J. W. Jeon, IEEE Trans. Circuits Syst. Video Technol., vol. 20, no. 15, (2010).
- [8] X. Sun, X. Mei, S. Jiao, M. Zhou and H. Wang, Editors, "Stereo Matching with Reliable Disparity Propagation", Proceedings of the 2011 International Conference on 3D Imaging, Modeling, Processing, Visualization and Transmission, Hangzhou, China, (2011) May, pp. 132-139.
- [9] H. Hirschmuller and D. Scharstein, IEEE Trans. Pattern Anal. Mach. Intell., vol. 31, no. 1582, (2009).
- [10] R. Zabih and J. Woodfill, Editors, "Non-parametric Local Transforms for Computing Visual Correspondence", Proceedings of the 3th European Conference on Computer Vision, Stockholm, Sweden, (1994) May, pp. 151-158.
- [11] X. Mei, X. Sun, M. Zhou, S. Jiao, H. Wang and X. Zhang, Editors, "On Building an Accurate Stereo Matching System on Graphics Hardware", Proceeding of the 13th International Conference on Computer Vision, Barcelona, Spain, (2011) November, pp. 467-474.
- [12] J. K. Tanskanen, T. Sihvo and J. Niittylahti, IEEE Trans. Circuits Syst. Video Technol., vol. 14, no. 1270, (2004).
- [13] Middlebury stereo datasets, http://vision.middlebury.edu/stereo/data/.
- [14] Wikimedia Commons, http:// http://commons.wikimedia.org/wiki/File:Lion\_(parallel)\_(163984575).jpg.

### Authors



**Hyeon-Sik Son** is a Ph.D. student in the School of Electrical Engineering & Computer Science, Kyungpook National University at Daegu in Korea. He received the B.S. degree and M.S. degree in Electrical Engineering & Computer Science in Kyungpook National University in 2010 and 2012 respectively.

His research interests are in the following areas: SoC (system on a chip), VLSI, hardware architecture, and stereo vision.



**Kyeong-ryeol Bae** is a Ph.D. student in the School of Electrical Engineering & Computer Science, Kyungpook National University at Daegu in Korea. He received the B.S. degree and M.S. degree in Electrical Engineering & Computer Science in Kyungpook National University in 2009 and 2011 respectively.

His research interests are in the following areas: SoC, computer architecture, and computer vision.



**Yong-Hwan Lee** is currently a professor in the School of Electronic Engineering, Kumoh National Institute of Technology at Gumi in Korea. He received the B.S. degree, M.S. degree and Ph.D. degree in Electronic Engineering in Yonsei University in 1993, 1995 and 1997 respectively. He spent four years as a Senior Researcher in Hynix Semiconductor Inc., and also spent two years as a Senior Researcher in Samsung Electronics.

His research interests are in the following areas: SoC, embedded system and software, and microprocessor architecture.



**Byungin Moon** is currently a professor in the School of Electronics Engineering, Kyungpook National University at Daegu in Korea. He received the B.S. degree and M.S. degree in Electronic Engineering in Yonsei University in 1995 and 1997 respectively. He received a Ph.D. degree in Electrical & Electronic Engineering in the same university in 2002. He spent two years as a Senior Researcher in Hynix Semiconductor Inc., and also worked as a Senior Researcher in Yonsei University for one year.

His research interests are in the following areas: SoC, computer architecture, and vision processor.

\* Corresponding Author: Byungin Moon.