Hardware implementation for Real-time Census 3D disparity map Using dynamic search range.

Jueng Hun Kim
Sungkyunkwan University
School of Information and Communication
Suwon, Korea
jhunkim@vada.skku.ac.kr

Chan Oh Park
Sungkyunkwan University
School of Information and Communication
Suwon, Korea
copark@vada.skku.ac.kr

Jun Dong Cho
Sungkyunkwan university
Department of Mobile system engineering
Suwon, Korea
jdcho@skku.edu

***Abstract*— in this paper, a new hardware implementation for the real time disparity map is presented. The real time disparity map includes the Dynamic Search Range Estimation, Disparity estimation, and Error Correction. Our demonstrated design flow shows an approach to implementation and hardware architecture of real-time disparity map estimation. This design is efficiently synthesized on Xilinx vertex 4 FPGA. The resulting hardware implementation is analyzed and simulated for system clock speed 100MHz to verify adequate performance. Since the algorithm is not so complicated to adapt real-time hardware design, we successfully made system with FPGA Prototype design.**

Keywords-component; Image filter, Image processing; Mopological operation; Stereo Vision; Computer vision.

I. INTRODUCTION

Recently, motivated by the demand of 3D movie and television, the quality of 3D stereovision has been improved dramatically. Thus, a number of algorithms have been proposed []. The algorithm can be classified into local and global method. Local methods define the disparity of a pixel based on the support window matching. Since these methods usually have less complexity, the real-time architectures are often implemented by the local methods. Global methods define the disparity of all the pixels as a global energy function. Nonetheless, finding the optimized global energy function requires too much complexity metric operation. Thereby, we concern on local methods only.

Local methods originally considered the comparison of various similarity measures []. Zabih have done studying the census transform and the rank transform to obtain an improved disparity map [2]. Kuhn extended Zabih`s work to modified census transform to improve hardware efficiency [8]. Chang studied the performance and speed cooperatively of similarity measures and color representation [17]. Recently, Chang (2010) extended his work to mini-census adaptive support weight to obtain high-performance real-time stereovision system [1]. Those works showed that the census-based color representation method achieved good performance among the local methods.

However, many of these algorithms are not enough to acquire high quality disparity map image. Furthermore, most of them are so slow that it is almost impossible to implement the real-time disparity map image acquiring system without support of massive dedicated hardware like General-Purpose Graphic Processing Unit such as CUDA. Thus, modified Census algorithm mainly based on the Chang`s paper [1] is developed to overcome these weaknesses of previous algorithms.

We propose a real-time stereo matching system for VGA size images using the hardware-friendly disparity estimation algorithm. In order to reduce the data access, we apply the data pipeline for input node. Furthermore, using the adaptive threshold from simulation, we propose a simple dynamic search range estimation to preprocess the disparity estimation in the real-time operation with two different kinds of search range (mask). As a result, our system processed the disparity map of 640 by 480 at average 45 fps.

The remainder of this paper is organized as follows. Section 2 introduces modified census algorithm and Dynamic search range estimation. Section 3 introduces hardware architectures with several implementation methods. The last section shows the experiment result and conclusion.

II. CENSUS TRANSFORM AND HAMMING DISTANCE

The census transform is based on block matching method [1]. The census transform compares the intensities of neighbor pixels with a support window with the center pixel. If a pixel intensity is greater than the center pixel intensity, the comparator will give '11', else if, a '00' bit. The hamming distance of their signature strings compares two blocks.

The example of the hamming distance is shown as Fig. 2. The hamming distance determines the number of bits that different between two bit streams. The result of hamming distance is the census value. The census value should be computed at every pixel with the searching range. To compute the correspondence, each census value is added with aggregating value. The pixel with the smallest value is selected for corresponding points.

Generally, aggregation is done by summing matching cost over square windows with constant disparity [10]. But the matching cost is calculated using squared difference. The square operation requires high computation. The other method of the preprocessing for disparity estimation is the search range estimation. The original disparity search range estimation method requires high computational complexity for high resolution images [21]. Thus we propose a dynamic search range estimation method for the preprocessing.

Figure1 shows the concept of the dynamic search range estimation. The histogram for search range is made by corresponding points on every line. The X axis of the histogram represents the initial disparity, and the Y axis of the histogram represents the number of the corresponding initial disparities. The initial disparity is the difference between the x coordinate of a pixel in left image and the x coordinate of the corresponding pixel in right image. Since the general image is composed with combinations of many objects, the disparities have close relations to each other. The defining threshold is shown in equation (1).

$$Threshold = \frac{Nmber\, of\, frequent\, initial\, disparity}{K} \quad (1)$$

The sparse disparity (i.e., with the small number of disparities) can be regarded as the noise and then the sparse disparities are eliminated by the threshold. The threshold in histogram to eliminate the sparse disparities is decided by the most frequently occurred initial disparity. This threshold checking will be started with the post dynamic search range estimation processing to give the reference information for full search range estimation. The result value will be reserved for the full search range estimation to check the entire disparity map.

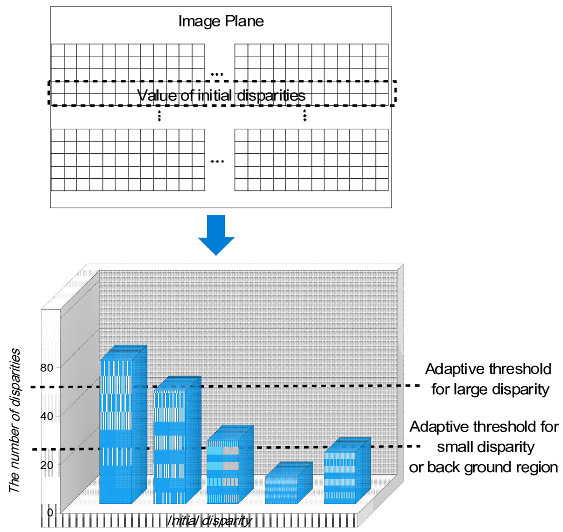

Figure 1. Search range estimation.

The proposed modified algorithm uses reference disparity map for full search range estimation. The concept is to Figureout the small disparity estimation and search range for unimportant pixel boundaries such as a huge same textured object, and back ground as shown in the Figure2.

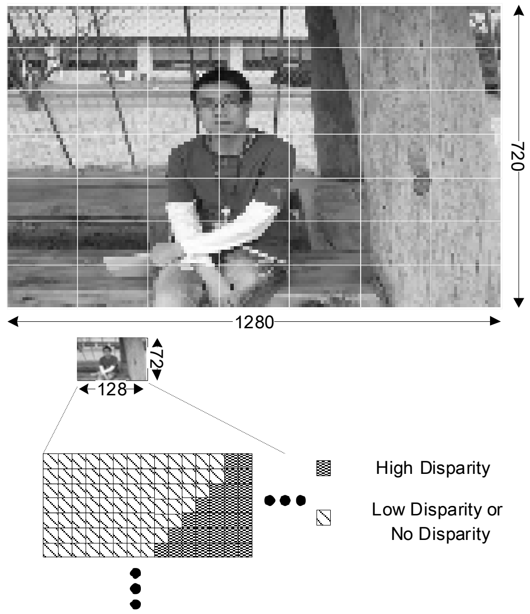

Figure 2. Different size of search range estimation.

There are two values: one for high disparity region, the other is low disparity or no disparity. Predetermine those values and use it as reference for full search range estimation. Since full search range estimation is the most time consuming module. Use of this method, the search range estimation time is much reduced.

After we obtain the final disparities based on the output of the census transform, we update the dynamic search range. In contrast to the earlier approaches, the updated dynamic search range is applied to the matching cost for evaluating the correspondence. If it is high disparity, the matching cost is set to the maximum value. In other case, the matching cost maintains the original search range. This decision flow is shown as in Figure 3.

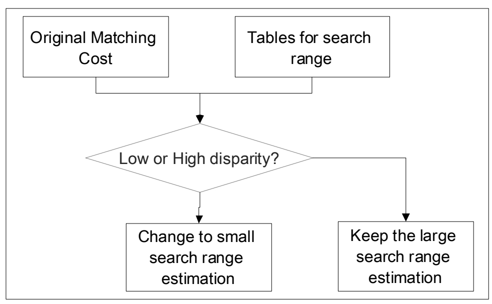

Figure 3. The method of the size of search range decision

We can see the histogram which is made from search range estimation shown in Figure 3. The higher possibility of matched area means that is has high histogram value, and it is likely to find matched window of right image window on the left image.

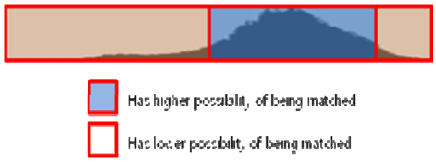

Figure 4. Matching histogram.

After finishing search range estimation, the full range estimation starts with calculating disparity with histogram simultaneously.

The first step is to acquire a pair of images from two cameras, which is located at different angles. We suppose that the two images are rectified when they are passed to proposed modified census algorithm hardware.

Histogram is needed as for the next step not to calculate the areas that have lower possibility of being matched. The algorithm scans all over the picture and accumulates histogram value. In Figure2, we can see the histogram, which is made from previous step. The colored area means that it has high histogram value and this means that it is likely to find matched window of right side of image window on the left side of image.

After finish all calculation, the error correction for noise reduction is processed. We have disparity map image of two images like Figure3. Darker pixels mean that the object there is closer to camera than brighter pixels.

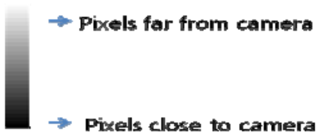

Figure 5. The distance concepts for bright to the dark color

III. Proposed Architecture

The Figure4 shows the overall hardware architecture for proposed census hardware architecture. We will describe this architecture with separate section search range estimation, Disparity map estimation, and Error correction.

A. Search range estimation

The Search range estimation hardware generates matching histogram for subsampling image pairs. Histogram value is embedded in these modules as like ram and the Disparity estimation module will read the histogram value. The Figure4 shows the architecture of search range estimation.

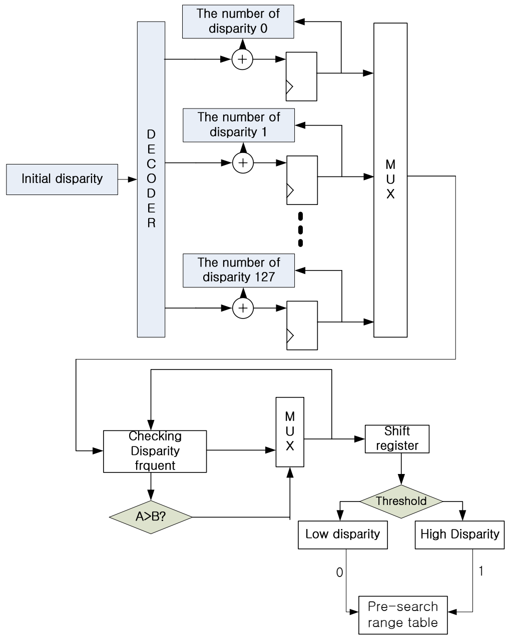

Figure 6. The hardware architecture of search range estimation.

B. Disparity estimation

The Figure5 shows the hardware architecture of Disparity estimation. This hardware calculates disparity of two images left side of image and right side of image. The result value will be stored rams called disparity image ram. Unlike search range estimation, this uses full size image. This architecture is the most time consuming section of the total time consumption for the overall hardware. However, by adapting several search range estimation and mini-map for dynamic search object can reduce computation time.

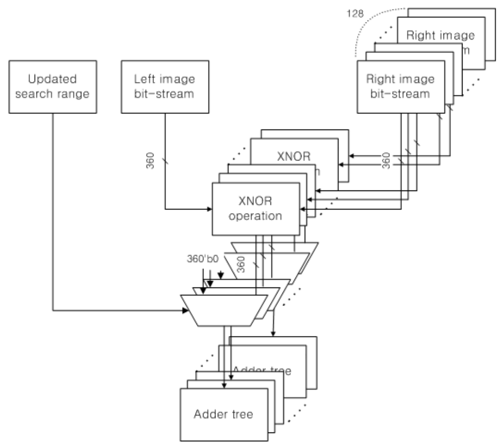

Figure 7. The hardware architecture of disparity estimation.

C. Error Correction

The function of error correction module is that reduce noise value in disparity map image which is made in wrong calculation from disparity estimation module. The noise reduction method is simple. Load four pixels around one pixel and test it whether it satisfies specific condition or not. In case it does not satisfy the condition, the value of result will be replaced the value of center pixel. The Error Correction hardware architecture is shown in Figure6.

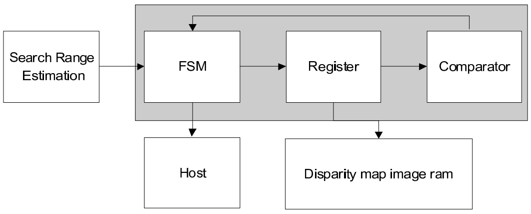

Figure 8. The hardware architecture of Error Correction.

IV. Experiment Result

Figure9 shows the disparity results of the proposed system with test images. We test the proposed system with the Middlebury stereo-pairs, which consist of Tsukuba, Venus, Teddy, and Cones [10].

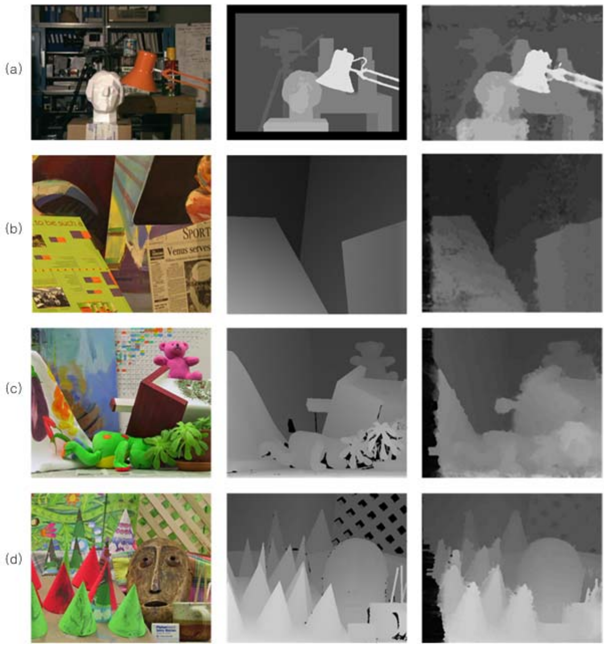

Figure 9. Disparity results of the proposed system. (Stero fair(left), Ground Depth(middle), Proposed algorithm(right))

Table 1 shows the proposed system performance.

TABLE I. Implementation Result

Xilinx Virtex 5 vlx330 prototype (1280x720 image)			
	# of LUTs	***# of BRAM***	***Max frq***
Census transform	11,006(5%)	165(57%)	232.342MHz
Hamming distance	2,677(1%)	N/A	389.105 MHz

Xilinx Virtex 5 vlx330 prototype (1280x720 image)			
	# of LUTs	***# of BRAM***	***Max frq***
Presearch Seach range	56,985(20%)	80(26%)	139.237 MHz
Full search range	39,187(16%)	N/A	120.32 MHz

TABLE II. Speed of Proposed Streo Vision System

Xilinx Virtex 5 vlx330 prototype (1280x720 image)		
MDE per second	***Search Range***	***Frames per second***
7077.9	128	28
4522	64	35
4233	Adapting Dynamic Search range	Average 30 Depends on the image.

V. Conclusion

In this paper, we proposed disparity estimation using dynamic search range for the real-time system. We have implemented an FPGA-based stereo matching system using census transform. Our modified census transform window can cover large area in the image with low computation complexity. Our modified census transform adapts the dynamic search range method for increasing speed of searching disparity time using resizable sparse mask operation. The pipelined register also helps to increase input node so that the hardware could be parallelized, and thus speed of computational time is increased. Our system occupied 296, 578 LUTs when it is implemented in Xilinx Virtex-5 vlx330 FPGAs.

Acknowledgment (Heading 5)

References

[1] N. Y. Chang, T. Tsai, B. Hsu, Y. Chen, and T. Chang, “Algorithm and Architecture of Disparity Estimation With Mini-Census Adaptive Support Weight,” IEEE Transaction on Circuits and system for Video Technology, Vol. 20, No. 6, 2010.

[2] R. Zabih and J. Woodfill, “Non-parametric Local Transform for Computing Visual Correspondence,” In Proceedings of the third European conference on Computer Vision, Vol. 2, pp. 150-158, 1994.

[3] J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.

[4] I. S. Jacobs and C. P. Bean, “Fine particles, thin films and exchange anisotropy,” in Magnetism, vol. III, G. T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271–350.

[5] K. Elissa, “Title of paper if known,” unpublished.

[6] R. Nicole, “Title of paper with only first word capitalized,” J. Name Stand. Abbrev., in press.

[7] Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, “Electron spectroscopy studies on magneto-optical media and plastic substrate interface,” IEEE Transl. J. Magn. Japan, vol. 2, pp. 740–741, August 1987 [Digests 9th Annual Conf. Magnetics Japan, p. 301, 1982].

[8] M. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.