Proceedings of

The International Conference on Intelligent Information Technology
(2INTECH 2015)
and
International Workshop on Advanced Computing and Multimedia Technology
(ACMT 2015)

October 21 - 23, 2015

Guam, USA

2INTECH 2015 and ACMT 2015

Table of Contents

Internet of Things Network System Using Virtual Access Points 1
Taekook Kim, Jai-Jin Jung and Eui-Jik Kim

TCP throughput gain in multi-radio wireless networks using ACK removal technique .. 6
Ayinebyona Eliab, Jongsun Park, Joo-Yub Lee and Jungmin So

A Fast Stereo Matching Algorithm and Its Hardware Architecture for Real-time Embedded Multimedia Systems 13
Kyeong-ryeol Bae, Byungin Moon

A Simplified Rectification Method for Efficient Hardware Implementation 20
Jongkil Hyun, Byungin Moon

A Study on Fast Partition Page Table Management for the DIMM Tree Architecture . 29
Young-Kyu Kim, Yong-Hwan Lee, Byungin Moon

A Design of System using Homomorphic Encryption for Multimedia Data Management 40
Hyun-Jong Cha, Ho-Kyung Yang, Jin-Mook Kim

Efficient Big Data Anonymization 55
Sung-Bong Jang, Young-Woong Ko

A Hybrid Approach to Nurse Re-rostering problem 72
Saangyong Uhmn,Young-Woong Ko, Kwang-Mo Lee and Jin Kim

A Grey Based Risk-minimizing Model Using Information of a Supply Chain 87
KyoungJong Park

How Different Connectivity Patterns of Individuals within an Organization Can Speed up Organizational Learning 92
Somayeh Koohborfardhaghighi, Dae Bum Lee, and Juntae Kim

Multimedia application to an extended public transportation network in South Korea: Optimal path search in a multimodal transit network 107
Yongshin Kang, Sekyoung Youm

Data Access Control Method for Multimedia Content Data Sharing and Security based on XMDR-DAI in Mobile Cloud Storage 114
Seok-Jae Moon, Jin-Mook Kim, Kye-Dong Jung, Jong-Yong Lee

A Fast Stereo Matching Algorithm and Its Hardware Architecture for Real-time Embedded Multimedia Systems

Kyeong-ryeol Bae, Byungin Moon

School of Electronics Engineering, Kyungpook National University, Daegu, Korea

bihmoon@knu.ac.kr

Abstract

Stereo matching is a procedure usually used to generate three-dimensional (3D) contents that are receiving attention in the field of multimedia. The most important considerations in stereo matching are highaccuracyof results and real-time performance. So, this paper proposes a vertical census transform with cost aggregation (VCT)to reduce hardware cost while maintaining high matching accuracy. The proposed census transformalgorithm uses one-dimensional vertical windows to calculate the partial cost of each pixel and generates the final cost by aggregating the partial costs across horizontal pixels at the same disparity. The experimental results show that the proposed algorithm reduces the hardware cost of stereo matching with higher accuracy compared with the conventional census transform algorithm. Thus, the proposed algorithm is suitable for real-time embedded multimedia systems.

Keywords:3D contents, stereo matching, census transform, vertical census transform, cost aggregation

1. Introduction

Multimedia is the use of text, graphics, animation, pictures, video, and sound to present information[1]. In recent years, three-dimensional (3D) contents generation and display technologies have been attracting great attention in the field of multimedia because these are used in various applications such as autonomous vehicles, intelligent robots,3D broadcasting systems and mobile devices.

Stereo matching,which is usually used to generate 3D contents, is a vision technique to calculatepixel disparitiesbetween the stereo image pair captured by a stereo camera and to acquire 3Dinformation by using the disparity. Among the stereo matching procedures, the most important step is to accurately calculate disparity from correspondence between stereo image pair. To find correspondence between pixels of stereo image pair, census transform algorithmthat is mostly used in stereo matching computes pixel correlation based on the relative ordering of pixel intensities within a window. The census transform is easy to implement and robust with respect to radiometric distortion[2]. Therefore, the census transform is not only used as a kind of local matching algorithm, but also used to calculate raw cost for global matching.

In the stereo matching for 3D contents generation, the most important considerations are accurate 3Dinformation and real-time performance. Besides, the optimization of hardware cost is also important for embedded multimedia system. There are many researches about hardware optimization and real-time performance of stereo matching based on the census transform. [3] proposed a census transform based stereo matching hardware architecture suitable for embedded real-time systems and [4,5] proposed modified census transform algorithms for hardware cost minimization. However, these hardware architectures required alarge amount of memory and theirhardware costs depend on the various factors such as image sizes, disparity search ranges (DSR) and window sizes.

Thus, this paper proposes a new stereo matching algorithm based on vertical census transform(VCT) with cost aggregation and a hardware architecture which helps the improvement of matching accuracy and reduces hardware complexity by using verticalcensus windows and cost aggregation. The remainder of this paper is organized as follows. Section 2 explains the conventional census transformalgorithm and proposes the VCT algorithm, and section 3 proposes a hardware architecture of the proposed VCT algorithm. Experimental results are presented in Section 4, and the conclusions are drawn in Section 5.

2. Vertical Census Transform with Cost Aggregation

The census transform algorithm introduced by Zabih and Woodfill is one of non-parametric local transforms [6]. The census transform algorithm is based on local intensity relationship between the center pixel and neighbor pixels within a certain census window. When $M\times N$ is selected as the census window size, the transform$T$ is defined as

$$T(x,y) = \otimes_{j=-N/2}^{N/2} \otimes_{i=-M/2}^{M/2} \xi\big(I(x,y), I(x+i, y+j)\big), \tag{1}$$

where $I(x,y)$ is the intensity of pixel (x,y), and the operator $\otimes$ denotes a bit-wise catenation. The relationship function ξ is defined as

$$\xi(p_1, p_2) = \begin{cases} 0, & if\ p_1 \le p_2, \\ 1, & if\ p_1 > p_2. \end{cases} \tag{2}$$

The hamming distance is calculated by a hamming function which determines the number of differences between two transformed census vectors. The hamming distance$C_d$is defined as

$$C_d(x,y) = Hamm\big(T_l(x,y), T_r(x-d,y)\big), \tag{3}$$

where *Hamm* is hamming function, and *d* is disparity, subscript *l* and *r* mean left and right respectively. After the hamming distance is computed, the final disparityD_{Final} is calculated using the winner-takes-all (WTA) method, shown as

$$D_{Final}(x,y) = argmin_{d\in R} C_d(x,y), \tag{4}$$

where *R* is a set of all possible disparity values.

For highly accurate matching results of the census transform algorithm, a large window is required because the matching accuracy of the algorithm strongly depends on the length of the census vector. However, the amount of computation and hardware cost is proportional to the length of the census vector.To overcome this problem, this paper proposes a VCT algorithm. It computes the hamming distance for vertical *N-1* pixels instead of all pixels within the census window (Figure 1) and aggregates hamming distances of adjacent windows in rowdirection (Figure 2).

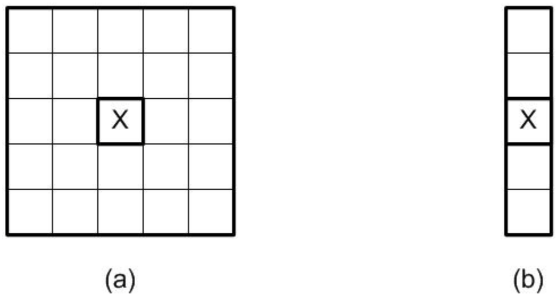


Figure 1**.** Window configurations. (a)Full census window.(b)Vertical census window.

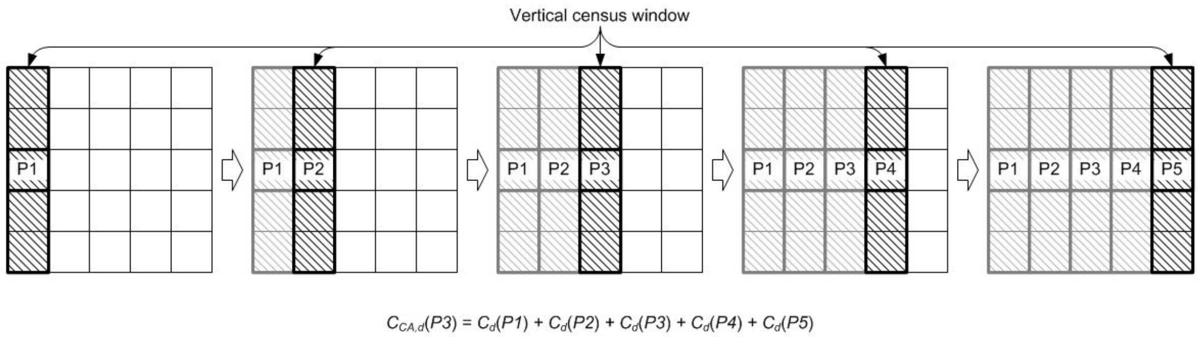


Figure 2. Cost aggregation of vertical census windows.

The transform $T$and the cost aggregated hamming distance $C_{CA}$of the proposed algorithm are defined as

$$T(x,y) = \otimes_{j=-N/2}^{N/2} \xi\big(I(x,y), I(x,y+j)\big) \tag{5}$$

and

$$C_{CA,d}(x,y) = \sum_{i=-M/2}^{M/2} C_d(x+i,y). \tag{6}$$

After the C_{CA} is computed, the final disparity of the proposed algorithm is calculated with C_{CA}, as shown in

$$D_{Final}(x,y) = argmin_{d \in R} C_{CA,d}(x,y). \tag{7}$$

Because of the modified transform procedure, the computational cost of the transform is reduced to $1/M$of that ofthe conventional census trasnform which uses the full window configuration. Also, during cost aggregation, box-filtering [7] can be adopted to eliminate ovellaped operations between adjacent census windows. For example, as shown in Figure 2, $C_d(P1)$, $C_d(P2)$, $C_d(P3)$, $C_d(P4)$ and $C_d(P5)$ are added for the $C_{CA,d}(P3)$. Also,$C_d(P2)$, $C_d(P3)$, $C_d(P4)$, $C_d(P5)$ and $C_d(P6)$ are added for the $C_{CA,d}(P4)$. However, by using the box-filtering, the $C_{CA,d}(P4)$ can be calculated as

$$C_{CA,d}(P4) = C_{CA,d}(P3) - C_d(P1) + C_d(P6). \tag{8}$$

3. Hardware Implementation

The computational complexity of stereo matching based on the census transform is easily calculated. When the size of window is $M \times N$, the image size is $W \times H$, and the DSR is R, the computational complexity are $O((M \times N\text{-}1) \times R \times W \times H)$ for computing the disparity in a frame. For an example, when the image size is 640×480, the window size is 3×3 and the DSR is 64, more than 5 billion operations per second are required for a processing rate of 30 fps. However, the proposed algorithm can reduce the computational complexity into $O((N$-

1)×*R*×*W*×*H*).For the real-time performance, this paper proposes a fully pipelined hardware architecture that cangenerate a final cost on every clock cycle by utilizingDSR-level parallelism (Figure 3). As shown in Figure3, input camera pixels are stored in Image Buffer and vertical *N* pixels are sent to the Transform. The Transform generates the vertical census vector by using vertical *N* pixels and the vertical census vectorsare stored in Left/Right vertical census vector buffers. For DSR-level parallelism, it has a Left vertical census vector buffer and *R* Right vertical census vector buffers. So the Vertical hamming function can computes*R* hamming distances between left and right vertical census vectors on every clock cycle. The computed hamming distances are storedinthe Delay buffersfor the cost aggregation and then final costs generated by cost aggregation are sent to WTAsimultaneously. The WTA calculates a disparity by finding the optimal cost and the disparity is refined by L/R cross check and Propagation.

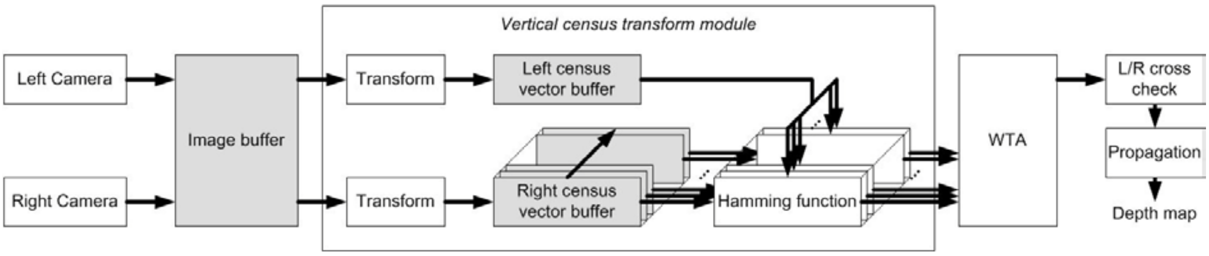


Figure 3. Stereo matching systemusingthe proposed vertical census transform with cost aggregation.

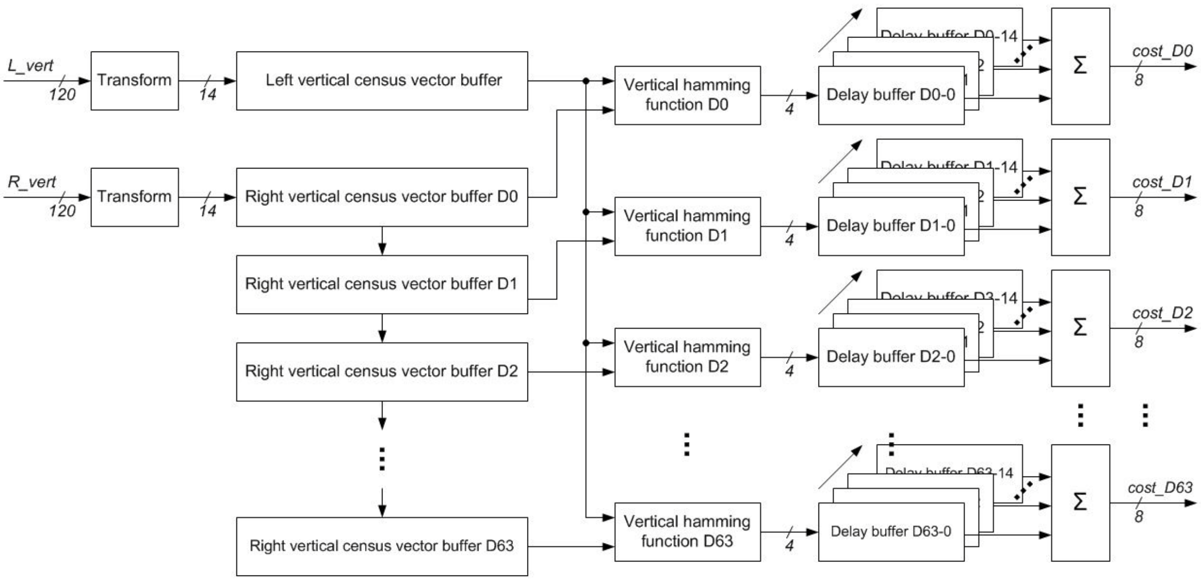


Figure 4. Proposed vertical census transform modulewith the widown size of 1×15 andthe DSR of 64.

4. Experimental Results

To evaluate the performance of the proposed algorithm, we modeled the stereo matching processors based on the VCT algorithm using C language and performed experiments with stereo image sets and their ground truth images. In addition, to show the hardware cost, stereo matching processors based on the proposed VCT algorithm and conventional census transform algorithm are implemented using HDL and verified on a FPGA-based platform. To investigate the matching accuracy of proposed algorithm, this paper performed experiments using four types of Middlebury stereo datasets (Figure 5).

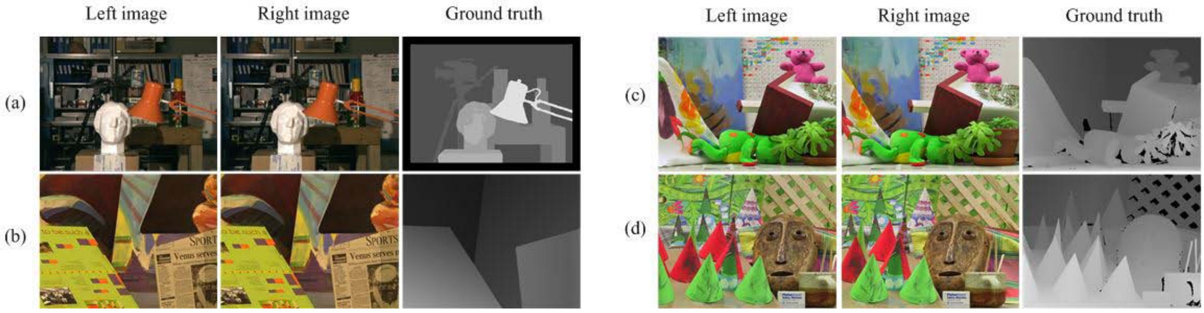


Figure 5. Middlebury stereo data sets used in the experiments. (a) Tsukuba.(b) Venus.(c) Teddy.(d) Cones.

First, this paper compared the matching accuracy of the proposed algorithm with that of the conventional census transform algorithm with various window sizes (Figure 6). In the case of the proposed algorithm, the window size means the size of one set of vertical windowsused for cost aggregation. As shown in Figure 6, the proposed algorithm outperforms the conventional census transform algorithm at every image. The major reason is that the final cost is generated by multiple windows. Each window has a center pixel repectively and this leads to blurring effect in results. So, many errors of conventional census transform algorithmcan be removed.

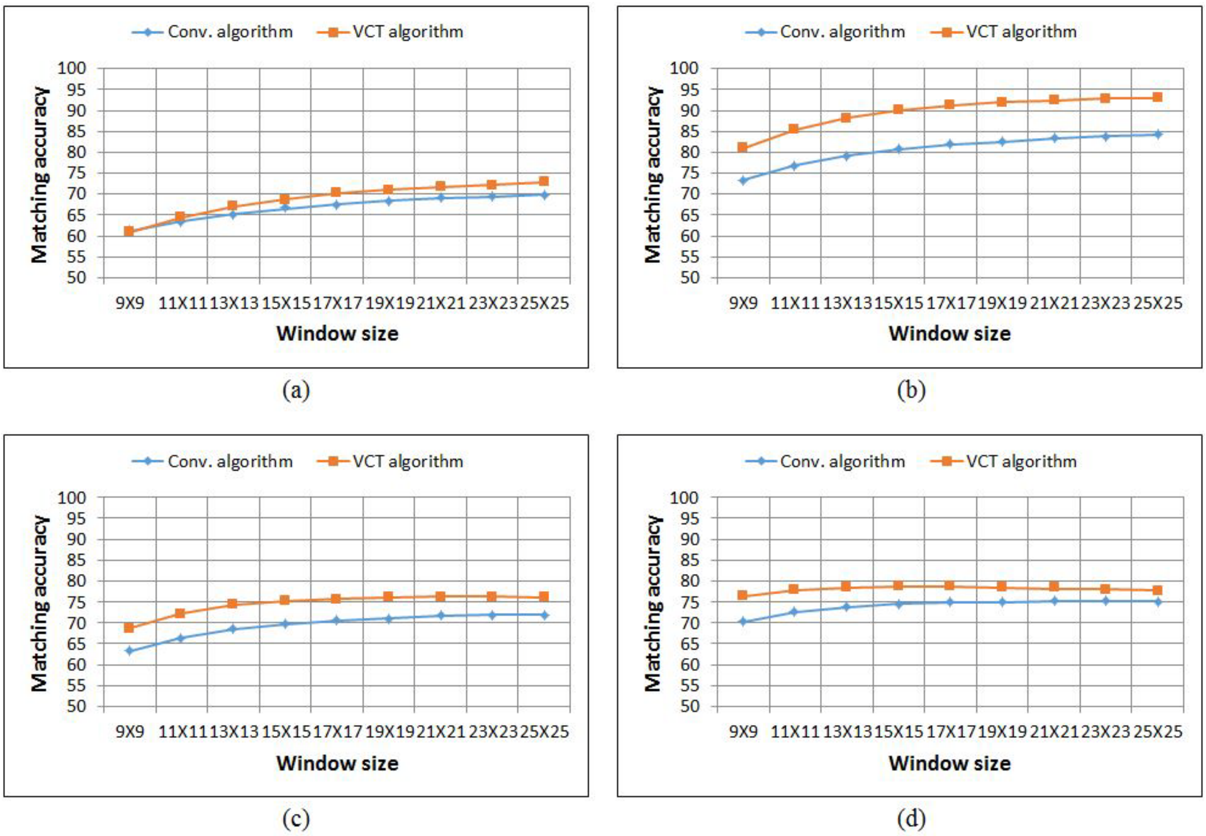
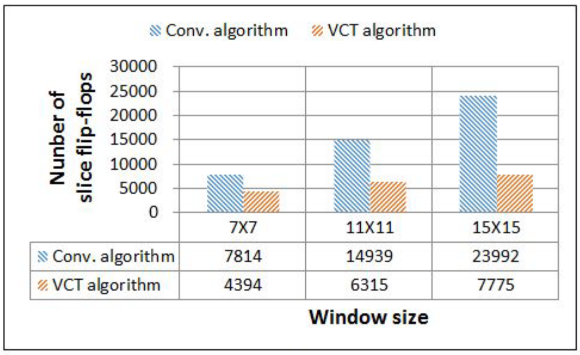
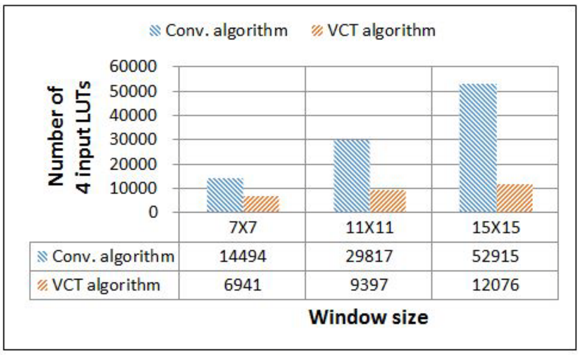


Figure 6. Comparison results of the matching accuracy between proposed algorithm and the conventional census transform algorithm with various window configurations. (a) Tsukuba.(b) Venus.(c) Teddy.(d) Cones.

Next, in order to demonstrate the advantages of proposed algorithm in terms of hardware cost, stereo matching processors based on the proposed VCT algorithm and conventional census transform algorithmare designedwith 3 different sizes of windows. The processors are designed using HDL and implemented on a Virtex-4 XC4VLX200-10 FPGA from Xilinx. All of these processors are able to process up to 325 640×480 disparity maps per second running at 100 MHz clock frequency. Thus, the processing speed comparison of these processors is meaningless and not discussed in the experiments.Figure 7 shows the comparison results of the FPGA hardware implementation. According to the experimental results, the proposed stereo matching processors use less than half hardware cost compared with the conventional census transform algorithm based stereo matching processors. In addition, as the window sizeincreases, the differences of hardware costs are rapidly increased. This is mainly because the hardware cost of stereo matching processors are proportional to the length of the census vector. In the conventional census transform algorithm, the length of census vector is proportional to $M \times N$. However, the length of census vector in the VCT algorithm is propotional to $1 \times N$.Thus, the proposed VCT algorithm easily supports large window sizes and is more suitable for real-time embedded multimedia systems.



(a)



(b)

Figure 7. Synthesis results of theFPGA hardware implementationof the proposed algorithm and conventional algorithm with various window configurations.(a)Number of slice flip flops.(b)Number of 4 input LUTs.

5. Conclusions

Stereo matching is the most widely used 3D contents generation method for various multimedia. Among stereo matching algorithms, the census transform is commonly used because of its characteristics. Many researches and products of stereo matching are introduced but many problems still remain about matching accuracy and hardware optimization. Thus, this paper presents VCT algorithm which helps the improvement of matching accuracy and reduces hardware complexity. The proposed algorithm reduces hardware costs dramatically by using the vertical window configuration instead of the full window configuration and improves matching accuray by using cost aggregation of hamming distances of adjacent vertical census windows in row direction. To evaluate the performance of the proposed method, this paper implemented the stereo matching processors based on the VCT algorithm and conventional census tranform algorithm using C and HDL. The results show that the proposed method generated more accurate results and consumed more than 50% fewer hardware costs than the conventional method. So the proposed method is suitable for real-time multimedia applications using the stereo matching.

Acknowledgements:
This research was supported by the MSIP (Ministry of Science, ICT and Future Planning), Korea, under the C-ITRC (Convergence Information Technology Research Center) (IITP-2015-H8601-15-1002) supervised by the IITP (Institute for Information & communications Technology Promotion).

Corresponding Author:
Associate professor Byungin Moon
School of Electronics Engineering
Kyungpook National University
Buk-gu, Daegu 41566, Korea
E-mail: bihmoon@knu.ac.kr

References

1. Najjar LJ. Multimedia Information and Learning. JI. Of Educational Multimedia and Hypermedia 1996;5(2):129-150
2. Banks J, Bennamoun M, Corke P. Non-Parametric Techniques for Fast and Robust Stereo Matching. Proceedings of IEEE TENCON’97 IEEE Region 10 Annual Conference. Speech and Image Technologies for Computing and Telecommunications, Brisbane, Qld., Austrailia Dec. 1997;2:365-368.
3. Jin S, Cho J, Pham XD, Lee KM, Park SK, Kim M, Jeon W. FPGA Design and Implementation of a Real-Time Stereo Vision System. Circuits and Systems for Video Technology, IEEE Transactions on Jan. 2010; 20(1):15-26.
4. Chang NYC, Tsai TH, Hsu BH, Chen YC, Chang TS. Algorithm and Architecture of Disparity Estimation With Mini-Census Adaptive Support Weight. Circuits and Systems for Video Technology, IEEE Transactions on Jun. 2010;20(6):792-805.
5. Fife WS, Archibald JK. Improved Census Transforms for Resource-Optimized Stereo Vision. Circuits and Systems for Video Technology, IEEE Transactions on Jan. 2013;23(1):60-73.
6. Zabih R, Woodfill J. Non-parametric Local Transforms for Computing Visual Correspondence. Proceedings of Third European Conference on Computer Vision (ECCV), Stockholm, Sweden May 1994;2:151-158.
7. McDonnell MJ. Box-Filtering Techniques. Computer Graphics and Image Processing 1981;17(1):65-70.