Optimizing a Virtual Re-Convergence System to Reduce Visual Fatigue in Stereoscopic Camera

Jae Gon KIM†, Nonmember and Jun-Dong CHO†(a), Member

SUMMARY In this paper, we propose an optimized virtual re-convergence system especially to reduce the visual fatigue caused by binocular stereoscopy. Our unique idea to reduce visual fatigue is to utilize the virtual re-convergence based on the optimized disparity-map that contains more depth information in the negative disparity area than in the positive area. Therefore, our system facilitates a unique search-range scheme, especially for negative disparity exploration. In addition, we used a dedicated method, using a so-called Global-Shift Value (GSV), which are the total shift values of each image in stereoscopy to converge a main object that can mostly affect visual fatigue. The experimental result, which is a subjective assessment by participants, shows that the proposed method makes stereoscopy significantly comfortable and attractive to view than existing methods.

key words: visual fatigue, virtual re-convergence, disparity estimation, stereoscopic camera

1. Introduction

As technology advances, we are able to watch TV and movies with three-dimensional (3D) image technology, permitting us to see depths similar to real life dimensions. Stereoscopic and understanding human visual system for stereoscopic viewing may be the key to understanding the binocular stereoscopy [1]. In computer vision, stereo images are captured similar to the way in which the human eye captures scenes.

One of the key topics in the stereoscopic camera is to reduce visual fatigue, whereas maintaining sufficient 3D reality. A conventional stereoscopic camera usually consists of two single-lens cameras. The main problem of conventional stereoscopic cameras is to have visual discomfort or fatigue whereas the 3D impact enhances (this means that the disparity of the main object in stereoscopy is increased.) In a stereoscopic image, the observer focuses to match the left and right images with converging eye movement, permitting each image to appear as stereoscopy by spatial perception of human nature. In this process, the human eye experiences conflict between accommodation and convergence that directly affects visual fatigue [2], [3]. Resolving the problem of visual fatigue or discomfort, whereas watching 3D displays, even for a short period of time, has been considerably researched in relation to stereoscopic displays [4].

Some key words for stereoscopic camera are as follow:

- **Accommodation**: Accommodation from distant to close objects for the camera is done by moving the lens farther from the film, since the focal length is fixed. In this way, the target object is focused on the imaging sensor to preserve a clear image view. There is an accommodation range in which the objects are seen sharply (focused). The difference between near and far point is expressed in units of length.
- **Vergence**: The inward or outward turning movements of the camera focus in convergence or divergence.
- **Convergence**: Convergence describes the eyes’ ability to divert eye optical axes horizontally in an inward or outward direction. In a stereoscopic camera, the lens of each camera projects a separate image of objects on each imaging sensor. This convergence is performed when the target object is in same position of each sensor.
- **Disparity**: Stereo disparity refers to the difference in image location of an object seen by the left and right camera. Each camera converges to the main object in each imaging sensor; it is seen as a single image, and all other objects, in front of or behind the point of convergence, can be seen to be double images with disparity.
- **Divergence**: Divergence has the opposite meaning to convergence. When objects in stereoscopy show disparity, these objects are shown as a diverged image.

Figure 1 illustrates situations capturing stereoscopy and the meaning of stereoscopic camera key words. In Fig. 1, the upper figure shows that the stereoscopic camera converges to the car, thus, the house is a diverged object showing disparity. Accommodation from distant (i.e., house) to close objects (i.e., car) for the camera is done by moving the lens farther from the film, since the focal length is fixed. The bottom figure shows that the stereoscopic camera is rotated to converge on the house, thus the car is a diverged object in this case. This movement of the stereoscopic camera is termed vergence. The converged object can be changed when a stereoscopic camera moves. However, most commercial stereoscopic cameras have a fixed position of lenses, and always converge to the same distance objects.

We can adopt the virtually control method of the optical axis of a stereoscopic camera to converge the target object in the divergence zone, instead of moving the camera lens mechanically [5]. This method requires a dense and accurate disparity-map that requires a high computational...
process. The disparity-map (also known as depth-map) is an image consisting of whole horizontal disparity information of stereoscopy. The horizontal disparity is defined as the distance between the row coordinates of the pixel locations of corresponding pixels in the left and right image of a rectified stereo pair. They first obtain a dense disparity-map, and then synthesize the virtual view images using the disparity-map. Figure 2 shows the result of the virtual view image changed from the original view. The shift value for virtual control is decided from the disparity-map of the original stereoscopy. In most cases, the shift value is set to the highest disparity in the disparity-map.

In this paper, we propose such a virtual control method and refer it to as virtual re-convergence, whose goal is to adjust the disparity-map to converge a main object automatically, whereas capturing stereoscopy. Convergence of the main object is essential to decrease visual discomfort or fatigue, because the human eye experiences conflict when convergence is not clear. Our virtual re-convergence uses an optimized disparity-map to determine the shift value of each image, considering effective disparity variance of stereoscopy in real-time. Our system facilitates a unique search-range scheme, especially for negative disparity exploration.

This remainder of this paper is organized as follows. Section 2 presents our disparity estimation system, which is a procedure to create a disparity-map, using geometry of a binocular viewing condition. Section 3 describes our decision process of a so-called “global-shift value” that is the shifted value of each image in stereoscopy to converge to the main object. Section 4 presents our experimental condition and result. Finally, Sect. 5 concludes our work.

2. Disparity Estimation Using Geometry of a Binocular Viewing Condition

2.1 Geometry of a Binocular Viewing Condition

Disparity is classified as the positive disparity and the negative disparity introduced in the Science journal [6]. In Fig. 3 (a), the intersections between left and right imaging sensor projections define the region that is “captured” by both cameras. Planes of constant disparities are indicated by horizontal lines for the positive (uncrossed) and the negative (crossed) disparities. Zero disparity plane (also known as zero parallax plane) is a converged domain of the stereoscopic camera. The zero disparity area is usually referred to as a comfortable zone or convergence zone [2]–[4].

In Fig. 3 (a), any objects in the green shaded section and the red shaded section are able to converge. The green shaded zone is always larger than the red shaded zone, when the zero disparity area exists. The ratio of the green shaded zone to the red shaded zone is still retained, even though the stereoscopic camera performs vergence [6].

As shown in Fig. 3, the positive disparity in the stereoscopic corresponds to uncrossed lines (b), whereas, the negative disparity in the stereoscopic corresponds to crossed lines (c). The negative disparity exhibits crosstalk that occurs between accommodations of each eye. Thus, negative disparity can incur more visual fatigue than positive disparity [2]–[4].

From Fig. 4, positive disparity shows a smaller disparity and object size than negative disparity, because the object in positive disparity is farther from a stereoscopic camera than the object in negative disparity. This result is related to the geometry of a binocular viewing condition, because the area of the negative disparity zone is smaller than the one of positive area in the diamond-shaped region.

The key substances mentioned above are as follow:

1. The negative disparity area is unfavorable for observers, due to causing a crosstalk between accommodations of each eye.
2. The objects in the negative disparity area show more disparity than the ones in the positive disparity area.
Therefore, we realized that virtual re-convergence is necessary, especially emphasizing the negative disparity area to reduce visual fatigue. An effective disparity-map on the negative disparity area is essential to perform virtual re-convergence. Therefore, in this paper, we propose a disparity estimation method to obtain an optimized disparity-map with such a unique search range determination for the virtual re-convergence application.

2.2 Disparity Estimation to Create Optimized Disparity-Map

Disparity analysis usually offers a fast estimate of the convergence error. However, obtaining a perfect disparity-map requires much computational power (referred to as the global approach). Therefore, to obtain an effective disparity-map, we resort to a limited search-range, which is calculated based on the estimated total disparities (referred to as the local approach). Census transform [7], [8] is a well-known local approach. However, it is non-trivial to estimate the total disparities without search range bounds. Thus, we should determine the specific bounds on the search-range to trade-off computation time and accuracy.

We already mentioned that an effective disparity-map on the negative area is essential to perform virtual re-convergence. That is, we need more accurate depth information on negative disparity; because negative disparity is regarded as having higher impact to reduce visual fatigue (the negative disparity can incur more visual fatigue than positive disparity.) Therefore, we considered using a larger search-range than existing systems. However, it is inefficient to use a large search-range due to excessive computational time. Therefore, we limit the search range by calculating the ratio between negative and positive disparity areas.

Figure 5 shows the process of acquiring the ratio of negative disparity distance to positive disparity distance using properties of an isosceles triangle. Figure 5 (a) shows the characteristics required for an archetypal disparity detector.
Angle \( \alpha \) is AOV (angle of view) of each lenses. Angle \( \beta \) is a grade of baseline to outside the AOV. The positive disparity distance is denoted as \( d' \) and the negative disparity distance is denoted as \( d'' \). From (a), we derived (b) using properties of an isosceles triangle. Therefore, the positive disparity distance is directly proportional to the full distance from baseline of the stereoscopic camera to the end of the overlapped accommodation line of each lens (c). The negative disparity distance is directly proportional to the distance from the baseline to the start of the overlapped accommodation line of each lens (d). Thus, we can get the ratio between \( d' \) and \( d'' \).

We used trigonometric functions, as in Eq. (1), to derive \( d' \): \( d'' \). Therefore, we obtained a derivation as in Eq. (2). Then, we defined the positive disparity distance proportion to \( \tan \beta \) and the negative disparity distance proportion to \( \tan(\beta - \alpha) \).

\[
\frac{d'}{\alpha} = \tan \beta, \quad \frac{d''}{\alpha} = \tan(\beta - \alpha)
\]

\[
\therefore d' : d'' = \tan \beta : \tan(\beta - \alpha)
\]

From Eq. (2), we set the ratio of the negative disparity search-range to the positive disparity search-range to be 3:1 through our stereoscopic camera properties with fixed angles, \( \alpha \) and \( \beta \). Therefore, our system facilitates a unique search-range scheme comprised of 96 negative pixels and 32 positive pixels whereas other systems are comprised of the same pixels in negative and positive ranges.

Let us review the basic idea briefly behind the census transform. The census transform is a form of non-parametric local transform to map the intensity values of the pixels within a square window to a bit string. The center pixel’s intensity value is replaced by the bit string composed of a set of Boolean comparisons, such as that in a square window. This comparison generates a bit-stream for each window (we shall illustrate an instance shortly).

The “window” from the second image with the smallest Hamming distance from the primary “window” is determined to be the best match for the pixel. As shown in Fig. 4, let us recall that the difference in the pixel “window” positions between the two camera frames is inversely related to the distance of an object from the camera. The greater the distance between matching “windows”, the closer an object is to the camera pair. Thus, we can build a disparity-map.

In this paper, we choose stereo matching with the census transform to estimate the disparity and obtain a disparity-map in real-time. Figure 6 describes the overall flow of our disparity estimation system.

As shown in Fig. 6, our disparity estimation system (to be especially used for virtual re-convergence) consists of three major steps: 1) census transform is executed, and the matching cost of each pixel is calculated; 2) for each pixel, the total cost of census matching executed on each YUV plane is calculated; 3) the best disparity can be obtained using the winner-takes-all (referred to as WTA hereafter) method. Adopting a WTA strategy, the disparity associated with the minimum cost value is selected at each pixel.

First, the census transform matches a support window of pixel intensities around a central pixel to several candidate matching windows in another image. This comparison generates a bit-stream for each window. Figure 7 illustrates an example with the support window size of a \( 7 \times 7 \) template. If a pixel’s intensity is greater than the central
Table 1 Comparison with existing systems (* Note: ‘N’ means negative area; ‘P’ means positive area).

| Implementations          | Image Size | Matching Method          | Windows Size | *Search-range | Rectification | Frame Rate |
|--------------------------|------------|--------------------------|--------------|---------------|---------------|------------|
| Proposed (FPGA)          | 640 x 480  | Census Transform         | 11 x 11      | N: 96, P: 32  | Implemented   | 60 fps     |
| Jiui’s FPGA [8]          | 640 x 480  | Census Transform         | 11 x 11      | N: 32, P: 32  | Implemented   | 230 fps    |
| Yang’s GPU [10]          | 256 x 256  | Sum of Square Differences (SSD) | Variable-window | N: 50, P: 50 | None     | 6 fps      |
| Darabsha’s FPGA [11]     | 256 x 360  | Local Weighted Phase-Correlation | N/A          | N: 20, P: 0   | Pre-processed | 30 fps     |
| Woodfill’s ASIC [12]     | 512 x 480  | Census Transform         | N/A          | N: 26, P: 26  | Implemented   | 200 fps    |

Fig. 7 Example census transform procedure.

The actual images usually include noise in occluded and textureless regions in the initial disparity estimation result [9]. Thus, we also devise a filter that computes to smooth the disparity value with the neighborhood disparities. Figure 8 shows the process of our post-processing, a so-called “disparity smoothing.” As illustrated in Fig. 8, disparity smoothing simply calculates the average on selected depth values that show small intensity variance compared to the central pixel. The results of disparity in uncertain areas and the disparity continuity are improved over those of conventional methods due to disparity smoothing. We also perform search-range estimation using pre-processing, before the census transform to enhance accuracy in disparity estimation.

Previous work for stereo matching uses a similar search-range [7]. That is, in the previous work, the search-range consists of the same bounds in negative and positive directions. Throughout our experimentation, we realized that we should set the search-range with different bounds on negative and positive directions. We compared our system, as shown in Table 1, to the existing disparity estimation systems with different search range schemes.

3. Decision Process of Global-Shift Value

We described obtaining an optimized disparity-map to be used for virtual re-convergence in the section above. Now, we are at the stage of selecting the main object to converge and to make a virtual view image based on the disparity value of the main object. We should decide the main object to draw a virtual view image. The shift value that we decided is termed “Global-Shift Value (GSV).” The GSV is total shift values of each image in stereoscopy to converge to a main object.

Figure 9 shows the overall process to obtain the GSV. In this figure, we can see the optimized disparity-map image and the result image after performing virtual re-convergence. In figure, we can determine that closest object in the optimized disparity-map is a fist. However, in our
result image after performing our virtual re-convergence, a body was converged due to our GSV decision process. The result that was produced by our algorithm was much more effective than the existing research (this will be shown in Sect. 4).

Our process uses the following five steps to obtain GSV.

**Step 1** (Create an optimized disparity-map)

This process was already explained in Sect. 2.

**Step 2** (Calculate a disparity-histogram)

We calculate the total count of each distinct disparity using the optimized disparity-map. Then, we stacked the data, which contains each disparity count, on the memory using a histogram. The histogram, which is a graphical representation, dictates a visual impression of the data distribution. The index of the disparity-histogram is a disparity value.

**Step 3** (Sort disparity values that exceed the threshold value)

The threshold is used for noise reduction. It also helps to select the GSV candidates. We determine that the threshold is usually set to 5% of the total number of pixels.

**Step 4** (Find maximum disparity value)

Next, we find the maximum disparity value among the set of disparity values stored in memory. The maximum disparity value is regarded as the major source of visual fatigue.

**Step 5** (Decide GSV)

The main idea of the GSV determination process is to select the closest object that occupies the largest proportion in stereoscopy. Recall that the closest object will have the maximum negative disparity value, as mentioned earlier.

We need to check if one of the disparity candidate values (i.e., GSV candidates[i]) coincides with the disparity value of the object located in the center of stereoscopy to finally determine the GSV. Then, the selected disparity value becomes GSV. Whereas, if there is no such disparity in GSV candidates, we select a GSV corresponding to GSV candidates[i] associated with the largest number of pixels. For example, if there are 500 pixels with GSV candidates[0], and 1000 pixels with GSV candidates[1], then we select the GSV candidates[1] as a final GSV. We include the pseudocode below to clarify this process.

**Decide GSV**

Input: GSV candidates
Output: GSV

1. Check disparity value of center object in stereoscopy
2. If ( disparity value of center object == GSV candidates[0] )
3. GSV = GSV candidates[0]
4. else If ( disparity value of center object == GSV candidates[1] )
5. GSV = GSV candidates[1]
6. else If ( disparity value of center object == GSV candidates[2] )
7. GSV = GSV candidates[2]
8. else if ( total pixel counts of GSV candidates[0] < total pixel counts of GSV candidates[1] )
9. If ( total pixel counts of GSV candidates[1] >= total pixel counts of GSV candidates[2] )
10. GSV = GSV candidates[1]
11. else if ( total pixel counts of GSV candidates[1] < total pixel counts of GSV candidates[2] )
12. GSV = GSV candidates[2]
13. else
14. GSV = GSV candidates[0]

The disparity value obtained from the selected closest object is GSV. Then, we shift the left and right image by GSV (this process is virtual re-convergence).
4. Experimental Conditions and Results

4.1 Experimental Set-Up

We implemented the entire virtual re-convergence system for experimentation. Figure 10 (a) shows the stereoscopic imaging camera that has the same specification with a highly adjustable jig. This stereo camera sends each image through the HDMI cable to the FPGA board as shown in Fig. 10 (b). The board contains the HDMI image capture board, Xilinx Virtex-5 FPGA chip, and JTAG. The HDMI image capture board consists of two input ports from the stereo camera and one output port to the 3D TV as shown in Fig. 10 (c). Our system runs at 30 frames per second.

4.2 Participants for Experiments and Evaluation Method

Twelve people voluntarily participated in this experimental research. Participants who were familiar with the stereoscopic used in this study were barred from the test to minimize the effect of prior experience. Participants were briefed on the experimental situation before being involved in the experiment.

We used three kinds of virtual re-convergence systems.

(a) (b) (c)

Fig. 10 Experimental setup of stereoscopic camera (a), FPGA board (b), and 3D TV (c).

The first is our system. The second and third were introduced by Park et al. (2004) [5] and Jin et al. (2010) [8], respectively. Jin’s system is not for virtual re-convergence. However, Jin’s system was modified as applying the re-convergence with their proposed search-range scheme for the experimental comparison purpose. In our experiments, four different stereoscopic instances were given for each virtual re-convergence situation. The first is a large object in a negative disparity zone; the second is a large object in a positive disparity zone. The third and fourth are small objects in negative and positive disparity zones, respectively. A large object implies an object that exceeds 30% of the total image size. A small object is less than 5% of a total image size. We used the DSCQS (double stimulus continuous quality scale) method for the evaluation and the ACR (absolute category rating) to minimize fatigability of participants [13]. The five-level Likert scale is (very uncomfortable) (uncomfortable) (a little uncomfortable) (comfortable) (very comfortable).

4.3 Experiment Result

Figure 11 shows our real-time virtual re-convergence image results. In the left-hand images, the large object (i.e., horse’s face) exhibiting considerable visual fatigue (above) is shifted for virtual re-convergence (below). However, in the right-hand images, the small objects (i.e., pen and hand) still have the “pop-out” effect. This implies that these small objects still remain in the negative disparity area. After performing our virtual re-convergence system, most objects have shifted to the positive disparity area. Thus, the total visual fatigue value is sharply decreased.

Figure 12 shows resulting images of various virtual re-convergence systems. The test images of (a) and (c) are provided by Middlebury. The original image shows the overlapped image of stereo pairs with non-correction. In the result, our proposed method exhibits the most sophisticated revision. In the case of (a), our proposed system and Park’s system converged on a white stone that is located in the center and is the closest object. However, Jin’s system converged on a yellow box that is behind a white stone, because Jin’s system lacks the negative disparity area information. In the case of (b), all systems showed a similar result. However, our system converged on the horse’s face much more accurately due to Step 5 described in our algorithm (in Sect. 3). In the case of (c), our system converged on the dolls located in the middle horizontal line. Disparities of small dolls located in the front line are of negligible quantity, as described by Step 3 in our algorithm. However, Park’s system converged on the small dolls, because Park’s system does not have the decision process of Step5 in our algorithm. Therefore, Park’s system converged to the closest object preferentially. Jin’s system is almost similar to our system, except for the search-range. Therefore, Jin’s system converged on the same objects, as our proposed system. However, the resulting image from Jin’s system showed a much more blurred image than our proposed system due.
Fig. 11 Comparison before and after performing our virtual re-convergence.

Fig. 12 Resulting images of various virtual re-convergence systems: (a) with the large object in the negative disparity zone, (b) with the large object in the positive disparity zone, and (c) with the small objects in the negative disparity zone.
to the lack of the negative disparity area information. In Fig. 12, the case of small objects in a positive disparity zone is omitted, because all virtual re-convergence systems obtained the same result.

To evaluate optimization of our system, a cost function is used. The cost function defines similarity measure in stereo vision [14]. In this paper, we use three matching method for test. These methods are based on window matching. To describe the cost functions, we denote various symbols as follow:

- \( L \): left image of stereoscopic \((q\) is its pixel\)
- \( R \): right image of stereoscopic \((p\) is its pixel\)
- \( d \): disparity.
- \( W \): the set of all pixel locations of the used window while centered at reference point \((0,0)\).
- \( R_i \): intensity at pixels in \( W + p_i \).
- \( L_{i+d} \): intensity at pixels in \( W + q_{i+d} \).

The first method is the census transform. The census transform cost function is defined as follow:

\[
C_{\text{census}}(i, d) = \sum_{(x,y)\in W+p} p(x,y,d)
\]

where, \( p(x,y,d) \) is zero when \( R_i \) is bigger than \( L_{i+d} \) and \( p(x,y,d) \) is one when \( R_i \) is smaller than \( L_{i+d} \).

The second method is the sum of square differences (SSD). The SSD cost function is defined as follow:

\[
C_{\text{SSD}}(i, d) = \sum_{(x,y)\in W+p} |(R_{xy} - R_i)^2 - (L_{i+d,y} - L_{i+d})^2|
\]

The third method is Local Weighted Phase-Correlation (LWPC). The LWPC cost function is defined as follow:

\[
C_{\text{LWPC}}(i, d) = 1.0 - \frac{\sum_{(x,y)\in W+p} (R_{xy} - R_i)(L_{i+d,y} - L_{i+d})}{\sqrt{\sum_{(x,y)\in W+p} (R_{xy} - R_i)^2 \sum_{(x,y)\in W+p} (L_{i+d,y} - L_{i+d})^2}}
\]

We evaluate stereo matching on rectified image pairs obtained from the Middlebury stereo website [15]. Table 2 presents resulting percentages of matches compared to ground truth images from the Middlebury [15] for the three cost functions. In the result, census shows significantly better than SSD, and yields best stability in terms of standard deviation among the three matching methods.

Table 3 shows the evaluation results of DSCQS compared to the existing systems. Note that when the large-sized objects are located in the negative disparity zone, our virtual re-convergence will change the result considerably. Whereas, when the small objects are located in the positive disparity zone, our virtual re-convergence will change the result slightly.

Figure 13 signifies the stability of our virtual re-convergence system. Note that radical distance change of a main target object increases visual fatigue greatly. Thus, we need to check how our GSV reacts to radical disparity changes of a main object. In our experiments, we observed that our GSV needs only two frames on the average to make the main object distance to be stabilized. Thus, there is no radical distance change of a main object.

5. Conclusion

In this paper, we presented an optimized virtual re-convergence system to reduce the visual fatigue effectively caused by stereoscopy. The idea is to converge on the main object in stereoscopy that can most affect visual fatigue. We used a unique search-range scheme, which is comprised of 96 pixels in negative and 32 pixels in positive regions, to determine the main object well. Our virtual re-convergence system exhibits better results than existing systems, using our unique search-range and GSV determi-
nation method. The results of subjective assessment experiments revealed the proposed method significantly reduces visual fatigue. Note that any virtual re-convergence system requires additional image information outside the original stereoscopic border, because the shift operation in the system may cause the converged image to escape from the original image boundary. Therefore, the captured images of the stereo camera should be larger than the standard image size (e.g., VGA, HD, and FHD). Our system can be practically applied to 3D camera, 3D TV, and 3D mobile applications.

References

[1] O. Schreer, P. Kauff, and T. Sikora, 3D Video Communication: Algorithms, concepts and real-time systems in human centered communication, John Wiley & Sons, 2005.
[2] K. Ukaia and P. Howarth, “Visual fatigue caused by viewing stereoscopic motion images: Background, theories, and observations,” Displays, vol.29, no.2, pp.106–116, 2008.
[3] O.M. Lambooij, W. IJsselsteijn, M. Fortuin, and I. Heynderickx, “Measuring visual discomfort associated with 3D displays,” Proc. SPIE, vol.7237, 72370K, 2009.
[4] M. Lambooij, W. IJsselsteijn, and I. Heynderickx, “Visual discomfort and visual fatigue of stereoscopic displays: A review,” J. imaging science and technology, vol.53, pp.030201-1–030201-14, 2009.
[5] J. Park, G. Um, C. Ahn, and C. Ahn, “Virtual control of optical axis of the 3DTV camera for reducing visual fatigue in stereoscopic 3DTV,” ETRI J., vol.26, no.6, pp.597–604, 2004.
[6] I. Ohzawa, G.C. Deangelis, and R.D. Freeman, “Stereoscopic depth discrimination in the visual cortex: Neurons ideally suited as disparity detectors,” Science, vol.249, pp.1037–1041, 1990.
[7] D. Scharstein and R. Szeliski, “A Taxonomy and evaluation of dense two-frame stereo correspondence algorithms,” Int. J. Comput. Vis., vol.47, no.1/2/3, pp.7–42, 2002.
[8] S. Jin, J. Cho, X.D. Pham, K.M. Lee, S.-K. Park, M. Kim, and J.W. Jeon, “FPGA design and implementation of a real-time stereo vision system,” IEEE Trans. Circuits Syst. Video Technol., vol.20, no.1, pp.15–26, 2010.
[9] S. Yoon, D.B. Min, and K. Sohn, “Fast dense stereo matching using adaptive window in hierarchical framework,” Advances in Visual Computing, vol.4292, pp.316–325, 2006.
[10] R. Yang and M. Pollefeys, “Multi-resolution real-time stereo on commodity graphics hardware,” Proc. IEEE Conference on Computer Vision and Pattern Recognition, pp.211–218, 2003.
[11] A. Darabiha, J. Rose, and W.J. MacLean, “Video-rate stereo depth measurement on programmable hardware,” Proc. 2003 IEEE Computer Society Conference on Computer Vision & Pattern Recognition, vol.1, pp.203–210, 2003.
[12] J.I. Woodfill, G. Gordon, and R. Buck, “Tyzx DeepSea high speed stereo vision system,” Proc. IEEE Comput. Soc. Workshop Real-Time 3-D Sensors Use Conf. Comput. Vision Pattern Recog., pp.41–46, Washington D.C., 2004.
[13] “Subjective assessment of stereoscopic television pictures,” ITU-R Recommendation BT.1438, 2000.
[14] H. Hirschmuller and D. Scharstein, “Evaluation of cost functions for stereo matching,” Proc. 2007 IEEE Computer Society Conference on Computer Vision & Pattern Recognition, pp.1–8, 2007.
[15] Homepage: vision.middlebury.edu/stereo.