芸術科学会論文誌 Vol. 11, No. 4, pp. 166 – 175

Practical Stereo Matching Approach for Stereoscopic Video and Its Application

Junichi Nishikata[1)] **Eiji Sugisaki**[2)]

1) Graduate School of Computer and Information Science, Hosei University [*1]
2) Digital Magic, Ltd.
junichi.nishikata @ gmail.com, jil11hsg @ gmail.com

Abstract
Stereoscopic contents, such as 3D movies, Digital Signage and a glasses-free 3D etc. have recently become more available to the general public. Meanwhile, in terms of labor cost, creating stereoscopic contents still costs more money because it requires several manual labor processes such as drawing a depth map. In the creation of stereoscopic contents, depth information plays an important role for achieving a comfortable product by controlling stereoscopic effect and the depth information can be obtained by stereo matching methods with a set of input images. In this paper, we introduce an innovative approach suitable for the creation of stereoscopic contents, which is combining a fast method and a high-accuracy method. Our approach is constructed from the feature of depth perception and results of our questionnaire survey for verifying cognition of binocular disparity on background domain in stereoscopic contents. We also introduce our application for obtaining depth information by using our approach. Our application is interactive and intuitive, thus users can easily adjust parameters and obtain the results visually. Finally, the processing speed is evaluated by speed measurement and comparison with existing stereo matching methods.

[*1] He has graduated Hosei Univ. in March 2012

1 Introduction

In recent years, stereoscopic contents have become common due to reducing the price of hardware such as 3DTV, 3D video games, and mobile phones supporting 3D functions. In fact, the demand for the stereoscopic content creation has been expanding. In the process of the creation for stereoscopic contents, however, the production cost is still higher than the others because the creators need a quantitative design with careful considerations for range of fusion and the visible range of comfortableness. In addition, a disparity needs to be adjusted based on the screen size [1] and several types of artifact are generated according to the filming environment [2]. Therefore, practical know-how and manual labors are required to create excellent stereoscopic contents. Thus, there is a strong demand for a qualitative improvement of the technologies for reducing the cost of its creation.

Depth information can be utilized for adjusting a sense of depth in the creation of stereoscopic contents, such as applying the effect of a depth of field, compositing CG objects, modeling a 3D object, the readjustment of illumination effect, and creating a free viewpoint image. These examples tell that the information plays an important role for increasing the quality of the stereoscopic contents.

Depth information can be obtained by stereo matching with a set of several view-points. Stereo matching method has been one of the most active research areas in computer vision, and several methods have been proposed [3] [4] [5].

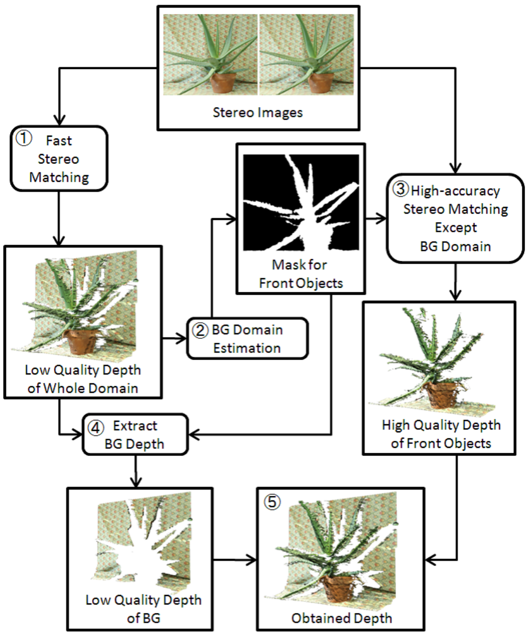

Figure. 1: Workflow for our approach.

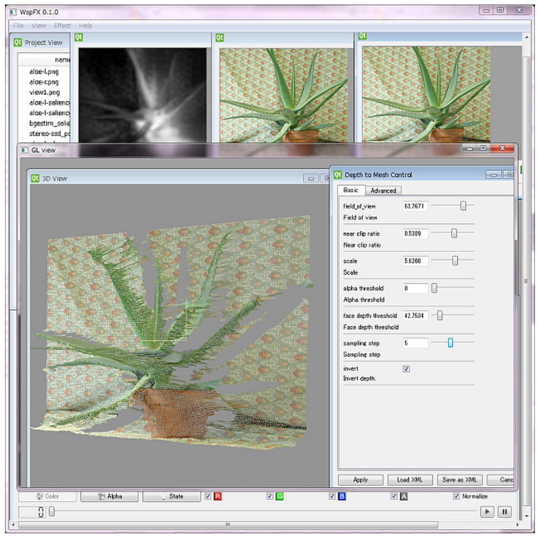

Figure. 2: Application overview.

There are fast methods in terms of calculation cost and high-accuracy methods in stereo matching methods, and they usually have opposite polarities. That is, an appropriate method must be selected depending on user's purpose and it is considered as a trade-off problem. In this paper, therefore, we propose an innovative approach suitable for the creation of stereoscopic contents by combining a fast method and a high-accuracy method.

Our approach is fundamentally according to the LOD (Level of Details). Figure 1 shows the workflow for our approach. First, rough depth information is obtained by a fast method. Second, we specify the domain that needs precise depth information based on the human's perceptual characteristic in stereoscopy and the features of attractive domain in the input. At this time, saliency map is applied to support this process. And then, we apply a high-accuracy method to the specified domain to obtain precise depth information.

Based on the proposed approach, we develop an application for computing depth information. Figure 2 shows the application overview. Users can control the obtained result interactively and intuitively and 3D structures are automatically generated based on the obtained depth information. The quality of the depth information obtained by our approach is indirectly evaluated by the questionnaire survey for evaluating stereoscopic video generated by shifting pixels based on an obtained depth. In addition, the processing speed of our approach is evaluated by the comparison with existing stereo matching methods.

2 Related Work

Based on the previous work and those ideas, we have proposed an efficient stereo matching approach and evaluated effectiveness and efficiency of our approach [6]. Actually, this paper inherits the concept and the purpose from the paper [6] and we develop an interactive and intuitive application for the practical use. We describe more details about related work in the following subsections.

2.1 Fast Methods

Dense stereo matching using image correlation is considered as one of the representative methods concerned with processing speed [3]. One of the methods, Window-based stereo matching, is to relate the data according to correlation in each local area. Moreover, a method based on image correlation is considered as the simplest stereo matching. In the local domain comparison between an image for left eye and an image for right eye, a high degree of similarity is associated with each other. In addition, SAD (Sum of Absolute Difference), SSD (Sum of Squared Difference), and NCC (Normalized Cross Correlation) are usually utilized for a correlation evaluation. These matching methods are suitable for real time processing, though there are several issues in the accuracy and errors.

2.2 High-Accuracy Methods

Several methods have been proposed for estimating sub-pixel disparities. There is a method for executing a correlation evaluation after upgrading a resolution of the target domain by image interpolation. In addition, correlation value functional fittings such as a parabola fitting and an equiangular line fitting are considered as the representatives of the functional fitting method [4]. Furthermore, there is a method estimating sub-pixel disparities by fitting POC (Phase-Only Correlation) values to a peak model function of POC [5]. The method based-on POC takes high computational cost but is more robust and more accurate enlarged resolutions than the others.

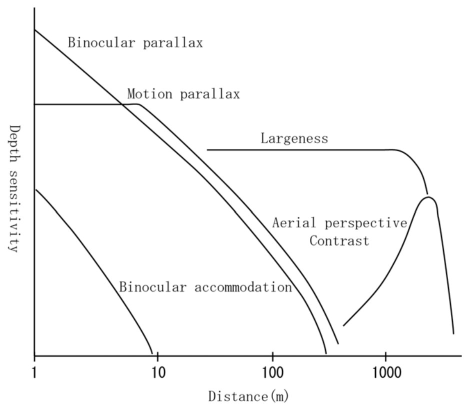

Figure. 3: Relationship between depth perception and the visual distance.

3 Characteristics of Stereoscopy and Attractive Domain

3.1 Characteristics of Depth Perception based on the Distance

There are two types of visual factors for perceiving a 3D object such as monocular depth perceptions (motion parallax, etc.) and binocular depth perceptions (binocular parallax, etc.) for depth perception. Because these factors become altered in accordance with visual distance from objects, it is assumable that the same phenomenon happens to stereoscopic contents. Nagata [7] has described the relationship between sensitivity for each case of depth perceptions and the visual distance from a subject to a target object. Figure 3 illustrates the relationships. Here, the depth sensitivity is defined as the ratio of the visual distance to depth discrimination threshold. This shows that an influence of binocular depth perceptions is reduced in proportion to the distance. On the other hand, monocular depth perceptions become more sensitive [7]. Figure 3 indicates that binocular depth perceptions predominate less than 10 meters and monocular depth perceptions are predominant more than 10 meters. In the result, while we need to more consider the binocular depth perceptions for the close-range medium such as a 3DTV, monocular depth perceptions need to be considered for the long-range medium such as a movie theater.

3.2 Attractive Domain

According to researches on the eye movements [8] [9] [10], humans tend to gaze at a moving object in the video. That is to say, the eye movements correspond to camera motions in most cases when watching the video. In addition, it is known that humans pay more attention to texts in the video. Additionally, related research [11] in stereoscopic contents has described that humans pay more attention to a human face. Furthermore, they observed that humans tend to be gazing more widely rather than gazing at a point if the content is stereoscopic. Therefore, in the case of a large distance between moving objects, texts, the main character (categorized as a front object), and the background domain, humans tend to gaze at front objects so the background domain is not focused. As a result, the feeling of video quality is less influenced by the quality of background domain.

4 Overview of Proposed Approach

4.1 Experiment for Verification of Our Approach

Based on characteristics described in section 3, we assume that the accuracy of background domain have little influence on binocular shape perception in stereoscopic video with front objects. For verifying this assumption, we have investigated by questionnaire how much the form accuracy of background domain affects the quality rating of a stereoscopic content. We suppose there are several front objects in a scene.

Figure 4 shows the example of investigation objectives. One is the flat background domain; the other is the detailed background domain. Specifically, we perform two investigations by five-grade evaluation to the 37 subjects.

1. watching stereoscopic video gazing at the front objects, (It could be said “watching normally”.)
2. watching stereoscopic video gazing in the background domain intentionally.

Table 1 shows the result of the questionnaire. In the result, while a half of trial subjects feel the clear difference in the case of gazing in the background domain intentionally, more than 80 percent of people do not notice the clear difference in the case of gazing at the front objects.

Table. 1: Results of the questionnaire:
- Cont. means that an observer watches two kinds of videos continuously,
- Sep. means that an observer can see an obvious switching between two videos such as showing black screen for a while.

Answer Choices	Answers			
	Attention to Front Objects		Particular Attention to BG Domain	
	Cont.	Sep.	Cont.	Sep.
Apparently No Difference	29 (78%)	21 (57%)	7 (19%)	6 (16%)
Little Difference	6 (16%)	10 (27%)	10 (28%)	13 (35%)
Slight Difference	1 (3%)	6 (16%)	4 (11%)	10 (27%)
Reasonable Difference	1 (3%)	0 (0%)	13 (36%)	5 (14%)
Absolute Difference	0 (0%)	0 (0%)	2 (6%)	3 (8%)

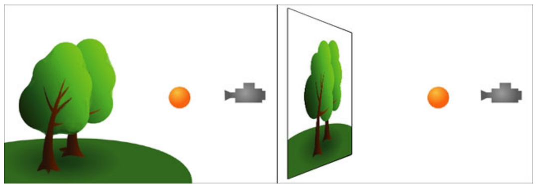

Figure. 4: Objective scenes: detailed background domain with 3D structure (left), flat background domain (right).

Here, we consider that the percentage of clear difference is the sum of “slight difference”, “reasonable difference”, and “absolute difference” in the Table 1. These facts support our assumption.

Consequently, detailed information of background domain does not affect much for the quality rating of stereoscopic contents. Thus, we assume that LOD of the background domain can be reduced for making the processing speed faster.

4.2 Execution Procedure

First, we apply a fast method to whole domain of an image and then apply a high accuracy method to the front domain. Here, front domain is defined as the domain other than background domain. The background domain and the front domain are determined automatically in our application.

Users can select from several correlations evaluation functions based on image correlation as the fast method in our application. Furthermore, in the high-accuracy method, users can also select from sub-pixel estimation methods: “by using scaling up input images”, “by function fitting of correlation value”, and “by function fitting of band limited POC”. That is to say, users can select one of these high-accuracy methods depending on their preference. The cost of our approach E can be determined by following equation.

$$E = C_A + C_C + C_B(1 - R_{bg}) \quad (1)$$

where C_A is the cost of a fast method, C_B is the cost of a high accuracy method, C_C is the cost of background domain estimation, R_{bg} is the ratio of the background domain in the input. This equation 1 shows that our approach has a feature that the processing speed becomes faster in proportion to the background domain.

5 Disparity Estimation by Fast Stereo Matching Method

The fast stereo matching method utilized in our approach is based on basic image correlation widely utilized in real-time in computer vision field. Users can select from SAD, SSD, NCC and others depending on their preference as an evaluation function for image correlation in our application. Because these evaluation functions vary a great deal in processing speed and accuracy, users can creatively utilize appropriate one according to the purpose.

In our stereo matching application based on image correlation, we first create a look-up table for a cumulative sum of correlation value in each disparity to both horizontal direction and vertical direction. And then, we compute a sum of correlation value in each correlation window. In this way, we make the processing speed of stereo-matching independent from the window size and achieve a high speed processing.

6 Estimation for Background Domain

An estimation for the background domain is required for determining the pixel to apply correspondence search by high-accuracy stereo matching method. However, there is a difficulty to numerically determine the background domain because it depends on several factors such as environment for screening a stereoscopic content. Therefore, it is extremely difficult to decide whether high-accuracy estimation is required or not.

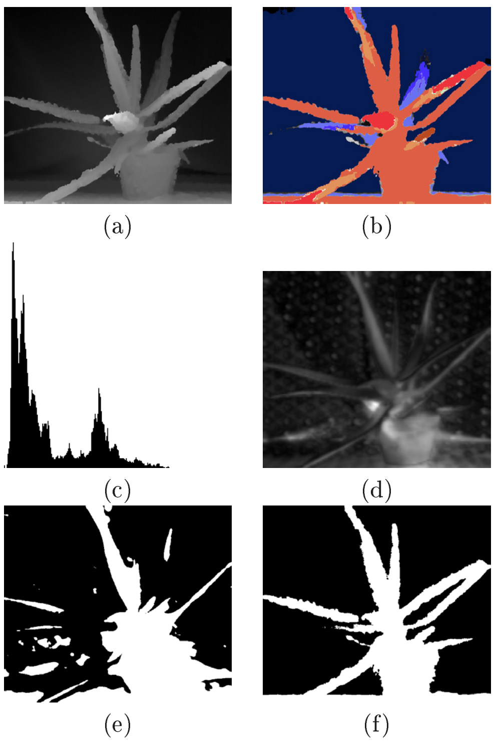

Figure. 5: (a) Depth map, (b) domain segmentation map, (c) histogram, (d) saliency map, (e) saliency-based mask, (f) mask for front objects.

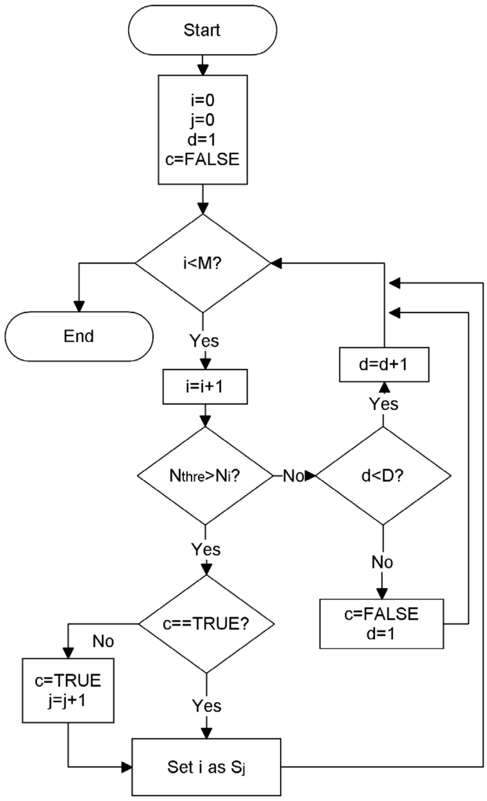

Figure. 6: Flowchart for depth segmentation based on histogram.

Thus, we achieve not only to implement an estimation for background domain but also to give users a domain segmentation of depth as an evaluation criteria for determining a boundary of the background domain.

By finding high density domain in a histogram shown in Figure 5 (c), the domain segmentation can be applied. The processes in detail are shown as follows.

1. Create a histogram from depth information
2. In the histogram, a continuous domain that satisfies condition $N_i > N_{thre}$ and $d < D$ as a domain S_j
3. Find S_{thre} from segments as a threshold for separating the whole domain into background domain and front domain.

The flowchart for process 2 is shown in Figure 6.

Where M is the number of depth levels in the histogram, N is the total number of pixels in the depth map, i is each depth level in the histogram, N_i is the number of pixels in i, N_{thre} is a threshold of N_i for determining whether the depth level has sufficient amount of pixels or not, S_j is each domain ID obtained in process 2, d is the number of steps between i and a depth level with the value more than N_{thre}, D is a threshold of d for determining whether i can be presence around S_j or not, S_{thre} is a threshold of domain ID for background domain.

In the background domain estimation process, we allow users to confine the candidate domains by using the saliency map for improving the reliability of estimated background domain. The saliency-based visual attention analysis is devised by Itti in 1998 [12]. The saliency effect can be obtained by features of intensities, colors, orientations from the multi-scaled image of input image. In addition to these features, we also utilize the optical flow feature for estimating the saliency. Figure 5 (d) shows the obtained saliency map from the input and Figure 5 (e) shows a processing mask generated by the saliency map.

7 Disparity Estimation by High-Accuracy Stereo Matching Method

We implement a stereo matching in sub-pixel disparity estimations for applying one of high-accuracy methods except the estimated background domain. In the stereo matching process, accurate depth information can be obtained by using sub-pixel disparity estimation at corresponding points. Details of implementation for the high-accuracy stereo matching in our application is shown in following sub-sections.

7.1 Image-Correlation with Upgraded Image Resolution

Stereo matching to the input with upgraded resolution can be considered as the simplest method for higher-resolution of disparity. This method is relatively easy to implement, but less efficient than the other methods, because it takes more time and costs more memory usage. However, since this approach can easily combine with other high-accuracy methods, there is a possibility to achieve much higher-resolution of disparity depending on combinations.

7.2 Functional Fitting for image correlation

In the image correlation method, the largest correlation value is determined to be a corresponding point. The corresponding point can also be determined by a fitting function using three of correlation values.

Basically the cost of functional fittings such as an equiangular line fitting and a parabola fitting are low, thus the functional fittings could be utilized for computing sub-pixel stereo matching fast.

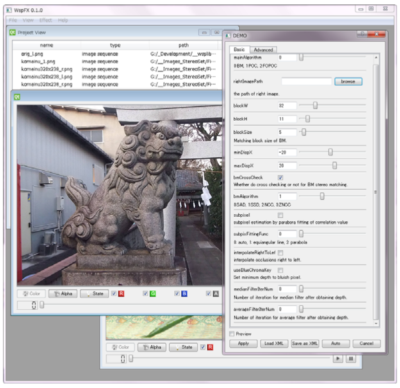

Figure. 7: GUI of our application.

7.3 Functional Fitting based on POC

In the method based on POC, we first extract only phase information from a frequency spectrum divided by amplitude and then calculate phase correlation. By using the calculated phase correlation, the peak model of POC function is obtained. Finally, the relations of rotation, scaling, and translation with a sub-pixel accuracy can be derived from the result of fitting the obtained peak model to the peak value around. POC peak functional fitting is more accurate, but offers higher computational cost, thus takes more processing time than functional fitting methods using image correlation such as SAD and SSD.

Because we only consider a horizontal disparity for stereo-matching, the fast POC method using 1D FFT is suitable for our application. Additionally in POC-based method, applying a band limitation makes the matching result more robust against noises. Therefore, we adopt a sub-pixel accuracy matching method based on 1D band-limited POC [13] as one of high-accuracy matching methods in our application.

8 Applications and System Evaluations

8.1 Application Overview

Figure 7 shows the main application GUI for our approach. GUI is developed based on Qt [14] and our application supports various file formats including animation sequence and achieve an animation playback with time-slider function.

- **Adaptive Dynamic Range Image**

  Dynamic range for disparity data depends on the input. Therefore we implement a compression format which cuts off unnecessary bits and then achieve most appropriate dynamic range. In this paper, we call this format as ADRI (Adaptive Dynamic Range Image).

$$R = D_{max} - D_{min} \quad (2)$$

Where D_{min} is the minimum of disparity data, D_{max} is the maximum of disparity data, and R is the necessary range. This format is quite effective. For example, if you have a binary image, the necessary range is only 1 bit, however it normally requires 8 bit range. On the other hand, ADRI can compress this as needed (in this case, just 1bit is the range.)

- **User-friendliness and Interactivity**

If you are an engineer, probably it is fun to configure various parameter settings. However, average users are confused or feel tiresome to find the best settings. Thus the parameters are separated into "basic" and "advance" in our application.

Additionally, we implement an automatic estimation function for the best settings and allow users to save their own settings into XML file. For achieving the interactive control, furthermore, we implement 3D viewer for checking disparity data that are shown as a 3D model or point clouds.

- **Plug-in**

For the advanced users' convenience, we design our application for easily developing users' own image processing filter as a plug-in. In fact, we prepare a SDK class provided for the users. The users can easily reflect the result of their own filter in our application by creating a new subclass of the prepared SDK plug-in interface. Users can add their own plug-in as they like.

In addition, we develop a Maya plug-in for transforming the disparity data into 3D structure and for editing its structure. This plug-in handles disparity data based automatic polygon generation according to the input which is an image format commonly-utilized or the ADRI format.

8.2 System Evaluation

In terms of practical use of our application, the process has been speeding up by parallelization with OpenMP.

We evaluate the processing speed for a fast method, a high-accuracy method, our approach respectively. Test objectives are (a) and (b) from Middlebury stereo datasets [15] (See Figure 8). (FYIG, the image in Figure 1 and 2 is from [16], those images can be downloaded from [17].) The resolution of these images are 450 × 375 pixels, and R_{bg} of these images are 0.234. Table 2 shows the experimental environment. The evaluations are performed in two cases. One is with an OpenMP parallelization, the other is without the parallelization. In each case, the number of trials is 100 times and its average of processing speed is utilized as an evaluation factor in each method. The disparity range of search in stereo matching process based on image correlation is between -51 pixels and -14 pixels in a horizontal direction. Experimental results are shown in Table 3, 4, 5. Table 5 shows respectively that Method A represents a fast method and Method B represents a high-accuracy method.

In these tables, following notations are utilized;

- the method based on image correlation with SAD as SAD BM
- the method based on image correlation with SSD as SSD BM
- the functional fitting method for image correlation with SAD using an equiangular line function as SAD ELF
- the functional fitting method for image correlation with SSD using a parabola function as SSD PF
- SAD BM with the resolution upgraded three times as SAD S3T
- SSD BM with the resolution upgraded three times as SSD S3T
- the method based on 1D band-limited POC as 1DBLPOC

We also evaluate the quality of depth information for the creation of stereoscopic content indirectly. Specifically, we first prepare the three of depth information: a fast stereo matching based on image correlation, a high-accuracy stereo matching based on 1D band-limited POC, and our approach. The evaluated objects are created by pixel sifting from -200 to 200 using these prepared depth information. We show the evaluated objects to subjects and then the quality of each evaluated objects are ranked. In this evaluation, we allow the subjects to give the same rank to each evaluated object. Figure 8 (d) ∼ (l) show the 3D models and free-viewpoint images respectively. We have 32 subjects in the evaluation and Table 6 shows each result of ranks.

Table. 2: Experimental environment.

Compiler	Visual C++ 2010 64 bit
OS	Windows 7 64 bit
CPU	Intel Core-i7 860 @2.8GHz
Memory	8.0 GB

Table. 3: Average processing time for fast methods.

Method	Block Size (pix)	Parallel	Time (ms)
SAD BM	7 × 7	YES	50
SAD BM	7 × 7	NO	114
SSD BM	7 × 7	YES	51
SSD BM	7 × 7	NO	114

Table. 4: Average processing time for high-accuracy methods.

Method	Block Size (pix)	Parallel	Time (ms)
SAD ELF	7 × 7	NO	180
SSD PF	7 × 7	NO	180
SAD S3T	13 × 13	YES	1413
SAD S3T	13 × 13	NO	4111
SSD S3T	13 × 13	YES	1404
SSD S3T	13 × 13	NO	4131
SSD 1DBLPOC	16 × 5	YES	454
SSD 1DBLPOC	16 × 5	NO	1621
SSD 1DBLPOC	32 × 11	YES	1511
SSD 1DBLPOC	32 × 11	NO	6493

Table. 5: Average processing time for our approach.

Method		Block Size (pix)	Parallel	Time (ms)
A	B			
SSD BM	1DBL POC	A: 7 × 7 B:16 × 5	YES	370
SSD BM	1DBL POC	A: 7 × 7 B:16 × 5	NO	1292
SSD BM	1DBL POC	A: 7 × 7 B:32 × 11	YES	1270
SSD BM	1DBL POC	A: 7 × 7 B:32 × 11	NO	4930

Table. 6: Result of quality evaluation.

Method	Vote Tally		
	1st	2nd	3rd
SSD BM	13	12	7
1DBLPOC	26	3	3
Proposed	23	8	1

9 Discussion and Future Work

9.1 Discussion

From the processing speed evaluation, the reduced time ratio is between 16% and 19% with an OpenMP parallelization and between 20% and 24% without the parallelization respectively though the background domain is about 23.4% in the whole image. The reason why the reduced ratio with the parallelization is smaller than the reduced ratio without the parallelization is considered that the parallelization cost itself affected the reduced ratio. In any case, the result from this evaluation proves that omitting the calculation cost for the background domain contributes to reduce the processing time in both cases, as expected.

In addition, the result from quality evaluation shows that the level of quality in our approach can be kept same as a high-accuracy approach. That is, there is no big difference between our approach and a high-accuracy approach in terms of human's perceived level though our approach is faster than a high-accuracy approach.

These results indicate that there is a high possibility that our approach is selected for a practical use because our approach can achieve decent quality with a high speed processing.

9.2 Future Work

We have described how we achieve a decent quality with a high speed processing for a stereo matching by omitting the calculation for the background domain in stereoscopic video which humans can hardly perceive visually. In addition, we have developed an interactive and user-friendly application for various levels of users.

However, there still are limitations and future work in our approach. Since it is formed from the hypothesis led from the conventional research for cognitive psychology [8] [9] [10] [11] and results from questionnaires in the previous sub-section, our approach can be applied to the stereoscopic contents only if there are several front objects in the scene. In other words, our approach is not reasonable and suitable for the stereoscopic contents that most of the domain is occupied by the background. Thus, we need to research several types of stereoscopic contents for finding domains beyond the reach of human's perceptions. If we can discover another domain such as the background domain we have handled in this paper, how we apply an automatic estimation to the discovered domain can be future work.

In addition, other combinations for stereo matching also need to be evaluated in terms of the processing speed and the quality.

Moreover, we have only evaluated the quality for the estimation of depth information for the purpose of a horizontal pixel-shifting. Therefore, we also need to evaluate our approach in accordance with the several intended use. There are several possible purposes such as an image processing effect based on the depth information, a pixel-shifting for free-directions, and a readjustment for illumination effect.

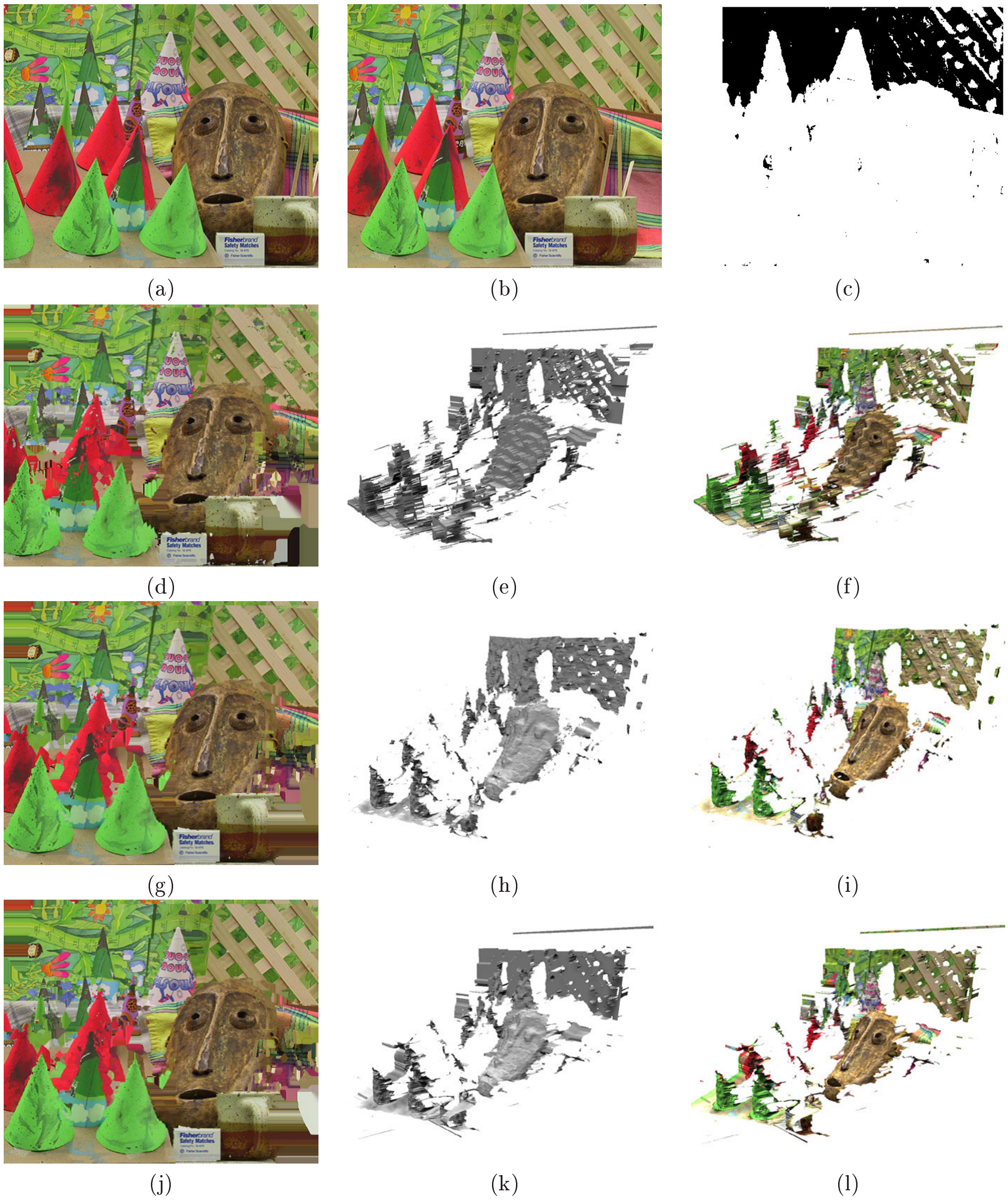

Figure. 8: Test objectives and results: (a) the input of left-eye image, (b) the input of right-eye image, (c) mask for front objects, (d) generated image by SSD BM, (e) obtained 3D model by SSD BM, (f) obtained 3D model with texture by SSD BM, (g) generated image by 1DBLPOC, (h) obtained 3D model by 1DBLPOC, (i) obtained 3D model with texture by 1DBLPOC, (j) generated image by our approach, (k) obtained 3D model by our approach, (l) obtained 3D model with texture by our approach.

References

[1] S.Kishi, S.H.Kim, T.Shibata, T.Kawai, "Scalable 3D image conversion and ergonomic evaluation. In Proc. SPIE," vol.6803, pp.68030F-68030F-9, Stereoscopic Displays and Applications XIX, January 2008.

[2] B.Mendiburu, "3D Movie Making: Stereoscopic Digital Cinema from Script to Screen," Focal Press, May 2009.

[3] D.Scharstein and R.Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms," International Journal of Computer Vision, 47 (1/2/3):7-42, April-June 2002. Microsoft Research Technical Report MSR-TR-2001-81, November 2001.

[4] M.Shimizu and M.Okutomi, "Precise Sub-Pixel Estimation on Area-Based Matching," Proc. 8th IEEE International Conference on Computer Vision (ICCV2001), vol.1, pp.90-97, Vancouver, Canada, July 2001.

[5] K.Takita, T.Aoki, Y.Sasaki, T.Higuchi and K.Kobayashi, "High-accuracy subpixel image registration based on phase-only correlation," IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences, Vol.E86-A, No.8, pp.1925-1934, August 2003.

[6] J.Nishikata, E.Sugisaki, "Efficient Stereo Matching Approach for Stereoscopic Video and Its Evaluation," NICOGRAPH International 2012, pp.153-160, July 2012.

[7] M.Nagata, "Visual Sensitivities to Cues for Depth Perception," Television, vol.31, no.8, pp649-655, 1977.

[8] O.Le Meur, P.Le Callet and D.Barba, "Predicting visual fixations on video based on low-level visual features," Vision Research 47, pp. 2483-2498, 2007.

[9] R.B.Goldstein, R.L.Woods and E.Peli, "Where people look when watching movies: Do all viewers look at the same place?," Computers in Biology and Medicine 37, pp. 957-964, 2007.

[10] J.E.Hoffman and B.Subramaniam, "The role of visual attention in saccadic eye movements," Perception & Psychophysics, Vol.57, No.6, pp.787-795, 1995

[11] J.Hakkinen, T.Kawai, J.Takatalo, R.Mitsuya and G.Nyman, "What do people look at when they watch stereoscopic movies?," SPIE, vol.7524, pp.75240E-1-10, 2010.

[12] L.Itti, C.Koch and E.Niebur, "A Model of Saliency-based Visual Attention for Rapid Scene Analysis," IEICE Transactions on Pattern Analysis and Machine Intelligence, Vol.20, No.11, pp.1254-1259, 1998.

[13] T.Shibahara, N.Numa, S.Nagashima, T.Aoki, H.Nakajima and K.Kobayashi, "A High-Accuracy Sub-Pixel Correspondence Technique Using 1D Phase-Only Correlation," IEICE Transactions on Information and Systems, vol.J91-D, pp.2243-2356, August 2008.

[14] Nokia Corporation, "Qt - A cross-platform application and UI framework," http://qt.nokia.com/

[15] D.Scharstein and R.Szeliski, "High-accuracy stereo depth maps using structured light," In IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2003), volume 1, pp.195-202, Madison, WI, June 2003.

[16] H.Hirschmuller and D.Scharstein, "Evaluation of cost functions for stereo matching," In IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2007), Minneapolis, MN, June 2007.

[17] Middlebury Stereo Datasets http://vision.middlebury.edu/stereo/data/

Junichi Nishikata

Junichi Nishikata is currently working as a software engineer in the game industry. He recieved a master's degree in information science from the Hosei University in 2012. His reseach interests include computer graphics, and image processing.

Eiji Sugisaki

Eiji Sugisaki is currently an Executive Producer and a Technical Director of Digital Magic Ltd., China. He has received a Ph.D degree from Waseda University, Japan and completed his time as a posdoc at Nanyang Technological University, Singapore. In addition, he used to be a visiting scholar at University of Illinois at Urbana Champaign, U.S. His research interests include non-photorealistic expressions, physics based hair dynamics, data-driven method to create a human motion, and image processing.