EMBODIMENT OF THE CORRECT DEPTH-CUE IN STEREOSCOPY

Kwang-Hoon Lee[1,2], Dong-Wook Kim[2], Yong-Moo Kwon[2], Nam-Ho Hur[3], Sung-Kyu Kim[2]*

[1]Department of Advanced Technology Fusion, Konkuk University, Seoul, Korea
[2]Imaging Media Research Center, KIST, Seoul, Korea
[3]Broadcasting & Telecommunications Media Research Division, ETRI, Daejeon, Korea
E-mail: geniuspb@kist.re.kr

Abstract

Pin-hole model has been widely used as a robust tool for easily understanding how to obtain a stereo image and how to present the depth-cue to an observer in stereoscopy. However, most of the processes to analyze depth cue in stereoscopy are performed that a stereo image is taken by camera model practically but depth cue of the image is analyzed by pin-hole model. Therefore, the result of depth cues by the process to be uncorrected. The reason of the uncorrected depth cue is led to the image distances of camera model due to variable focused object distances, and it makes a depth distortion. In this paper, we tried to show the contradiction such as occurring depth distortion in the process which the pin-hole model is used to analyze depth cue despite practical camera model is used in stereoscopy, and we presents the method to overcome the contradiction.

Keywords: stereoscopy, correct depth cue, depth distortion, focused object distance, image distance.

Introduction

Stereoscopy is well known to representative method based on the principle of binocular disparity to embody depth cues from stereogram which combined stereo images to an observer.[1] The process of depth recognition in stereoscopy is performed as following that object light is projected into the left and the right eye of the observer then the light make the object images having the inter-distance on the each retina. And then, the observer can recognize different depth-cue from their images formed on the each retina having disparity which was due to the distance of object. Disparity has charge of depth-cue of the stereo image and its value depends on the size of image on the each retina and inter-distance of eyes. Therefore, we can convert this stereoscopic process to using two light boxes system instead of human visual system (HVS). Pin-hole model is one of the systems based on geometrical analysis how to qualitative describe stereoscopic phenomena in the process. As you known that this model has a simple construction such as consists of a plane having pin-hole about 30 um by diffraction limit and detector plane forming object image. And it had advantages that easy to describe how to obtain a stereo image and how to present the depth-cue in stereoscopy. However, pin-hole model is not used practically in the process because it has long exposure time by very small size of hole to form the object's image on the detector such as a film or CCD. Camera model is the other representative system to convert HVS. This model used practically and frequently in the process more than pin-hole model because camera model has a lens system having a function of condensing light and forming clear image to the detector faster than pin-hole model. Therefore, the camera model has been used in most stereoscopic circumstances such as mobile, official, theater and so on. Let us see the viewpoint what type is adequate tool for analyzing correct depth-cue in the process when we have to select one in both types. If we take camera model to analyze depth-cue in the process, a property which carried out by this model will be different compared to pin-hole model was selected.[2] To put it more concretely that camera model has a parameter called the image distance and this parameter must be considered in the process, because the image distance is affected by focused object's distance (FOD). Image distance has meaning of inter-distance between representative nodal point of the lens of camera and the detector plane of camera when the object image was formed on the detector.[3,4] And it is defined as a function which related with the object distance in object's space and the focal length of the lens. So image distance let causes two phenomena that the field of view (FOV) will be changed relatively and the size of formed image on the detector to be either magnified or demagnified simultaneously. Furthermore, these phenomena lead to different disparity for the same object. In contrast, pin-hole model has fixed image distance whatever object is focused. Therefore, a value of disparity which above mentioned property will be different by both models when the same object is used in the process. If camera system is selected to obtain a stereo image, the method for analyzing depth-cue must be performed by camera model. However, a considerable number of studies have been reported that most of the processes to analyze depth cue in stereoscopy are performed as following that stereo image is obtained by camera system practically but depth cue from the stereo image is analyzed by pin-hole model[5,6,7]. It is clear that the result of depth cues by the reported process to be uncorrected. The reason of the uncorrected depth cue is caused by image distances in camera system due to variable FODs, and it makes a depth distortion. For providing it to an observer that correct depth information without depth distortion from the stereo-image taken by camera system, we need to a qualitative analysis on camera model with a relation between a value of disparity and a position of object where is focused.

In this paper, we tried to shows that the contradiction as a mentioned which depth distortion has been occurred in depth analysis by the reported existing method. And we proposed a new method to takes correct disparity and depth-cue by using a ratio of magnification of the size of stereo image which was taken by camera system. Finally, we had a significant result having corrected depth-cue from the compensated disparity in reconstructed stereo image by our method.

Analysis and Results

1. Qualitative analysis of a stereo image acquisition

1-1. Stereoscopy by the pin-hole model

The process taking a stereo image by pin-hole system is performed as following that. The planes of the pin-hole in the pin-hole boxes are aligned parallel to x axis in global coordinates and the inter-distance of centers of each the hole had set up as t. The origin of global coordinates is located in center of the inter distance and the objects in object space are located z axis increasingly. The normal distance between the hole and the plane of detector is fixed as f and the size of each hole is satisfied ray optics having diffraction limit about 30um generally. The size of the detector in the pin-hole box is w. Fig. 1 shows the geometrical scheme for obtaining the image coordinates for the object images which were formed on the detectors in each pin-hole box. The object image is projected on the detector as a point through the hole. The FOV in pin-hole system has fixed value wherever object distances are located in object space because the normal distance in pin-hole box is not affected by object's location. In Fig. 1, the bold lines show the projected object's rays from the point of object to the detector through the holes. And the dot lines on the background show the FOV of the pin-hole system. The object is located in x, z direction as $(0, z)$ and the coordinates of center of the detectors are $(-t/2, f), (t/2, f)$ in global coordinates. The denotations $w_{L,R}$, dis have meaning which as the image coordinates and the disparity of the object image on the detectors in the left and right of pin-hole boxes. The equation (1) shows the numerical solutions to extract the image coordinates taken by pin-hole model. And the equation (2) shows their disparity which is obtained by the equation (1).

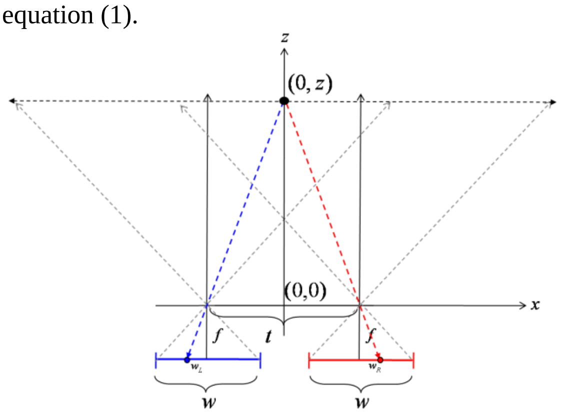

Fig. 1. The scheme to extract the coordinates of the object image on the detectors in pin-hole model

$$w_L = \frac{f}{z}\left(x + \frac{t}{2}\right) \;,\; w_R = \frac{f}{z}\left(x - \frac{t}{2}\right) \qquad \cdots\cdots\cdots (1)$$

$$dis \equiv (w_R - w_L) = \frac{-f\,t}{z} \;,\; \left\{\text{if } f < 0,\ dis = \frac{f\,t}{z}\right\} \cdots\cdots\cdots (2)$$

1-2. Characteristics of the camera model

The process obtaining a stereo image by the camera system is performed as following that. In this model, it has same circumstances in the pin-hole model such as camera alignment, inter-distance of cameras, size of detector, the origin in global coordinates and the location of object. The focal length of the lens in the both cameras is f. However, the image distance such as normal distance in case of the pin-hole box, will be different by object distances because it had a function of the object's distances based on a rule of the lens maker equation. That is a considerable point of difference between camera system and pin-hole system. The equation (3) show a calculation of the variances of image distances related to FODs. And denotes of f', s_0, f has meaning such as the image distances in the camera model, the FOD and the focal length of the lens. Camera system used in this paper is designed by Gaussian optics.

$$f' = f\left(\frac{s_o}{s_o - f}\right) \qquad \cdots\cdots\cdots (3)$$

If we want to make the camera system having same geometry as the pin-hole system, the FOD must be infinity, namely the image distance in camera system must be equal to the normal distance in pin-hole system. The Fig. 2 shows a scheme of the camera system having same geometric structure as the pin-hole system. In this case, the camera system was designed having a depth of field having lager than 22 of the f-number because the image blurring which is caused by shallow depth of field was not considered in the process.

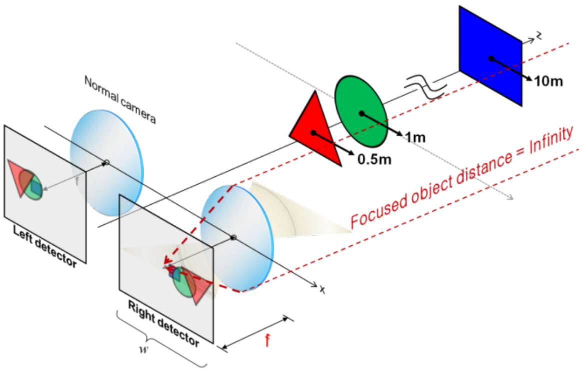

Fig. 2. The scheme shows the camera model having same structure geometrically as the pin-hole model.

In the camera model, the image distance will be increased when the FOD has not infinity. As this matter, it turns out that not only the FOV will be narrowed but also the size of image on the detector to be magnified. Therefore, the image coordinates on the detector will be different compared to the case of the pin-hole model although the same object was projected into both systems. Fig. 3 shows the scheme to obtain a stereo image by using the camera system when the FOD was not infinity.

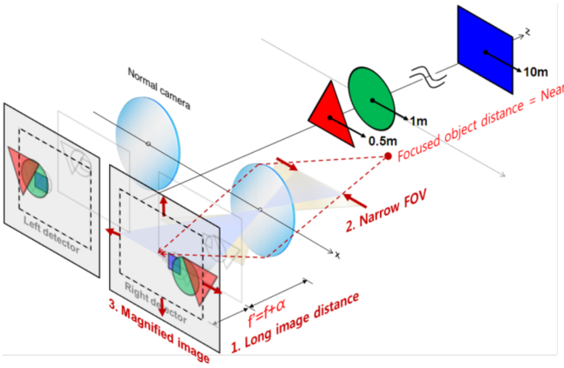

Fig. 3. The scheme to obtain a stereo image using the camera system when the FOD was not infinity

For this reason, the sizes of the formed object image and their image coordinates on the detector are magnified due to only variable FODs even if the object would be used same thing. The denotation as α present the values which increased the image distances due to variable FODs. The Fig. 4 shows that the scheme to presents the coordinate of the formed image on the detector and the position of the detector plane which having increased image distance when the FOD has not infinity. The object is located on $(0, z)$ in global coordinate. And the thin dot line presents the location of the detectors when the FOD had infinity and the bold line presents the location of it when the FOD had not infinity. The bold dot lines present the projection rays of object image into the detectors through the lens. And thin dot lines in background present FOV of the camera. As we can see Fig. 4, the FOV became narrowed comparing to the situation in Fig. 1. The Eq. (4) and (5) presents numerical solution to obtain the object image coordinates and its disparity when the FOD had not infinity. The denotation of $w'_{L,R}$ in Eq. (4) presents each image coordinates on the left and right detectors which has increased image distance. And the denotation dis' in Eq. (5) presents their disparity for the equation (4).

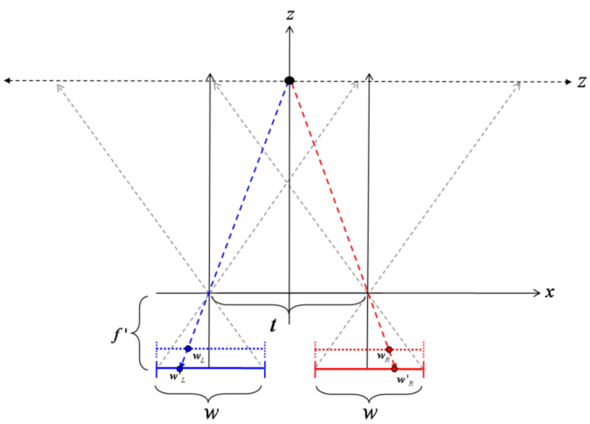

Fig. 4. The scheme to presents the coordinates of the formed image on the detector and the position of the detector which has increased image distance when the FOD has not infinity.

$$w'_L = \frac{f'}{z}\left(x + \frac{t}{2}\right) \quad , \quad w'_R = \frac{f'}{z}\left(x - \frac{t}{2}\right) \qquad \cdots\cdots\cdots (4)$$

$$dis' \equiv (w'_R - w'_L) = \frac{-f't}{z} \quad , \quad \left\{ \text{if } \ f < 0,\ dis' = \frac{f't}{z} \right\} \cdots\cdots (5)$$

2. Quantitative analysis of depth cue in displayed stereo image

In this paper, orthostereoscopy have been used as a method to embody the depth cue having a real scale for the object distance when the stereo image was displayed on the screen. In the process, the constraints were used to satisfy orthostereoscopy and they had having a same value such as the inter-distance of images, detectors and eyes of observer. And the ratio of image magnification is defined by a function which is derived the focal length of the lens over the viewing distance between an observer and the screen.

2-1 Analysis on the contradiction as a depth distortion

A stereo image on the detector obtained by camera system is transformed to a stereogram. And the stereogram would be either projected or displayed having magnifying ratio by user's circumstances such as mobile, office and theater. To embody the depth cue having a real scale to an observer, stereogram must be satisfied the constraints as we above mentioned. The contradiction is announced in introduction, it has means the results of the process which is performed as following that stereo image is taken by camera system but their depth cue is analyzed by pin-hole model. Namely, the result by this contradiction in the stereoscopic process will cause uncorrected depth-cue. The scheme in Fig. 6 shows that's why the contradiction is occurred. If each the image distance would be adequately considered to the process, the both image coordinates on the detectors as $w'_{L,R}$ and $w_{L,R}$ will have the same disparity. However, in the reported studies, the process of depth-cue analysis is performed by the fixed disparity in spite of increasing the image distance due to FOD. Hence, to take the correct disparity at stereo image, we must have considering the increased image distance which related to the FOD. In other word, the disparity of $w'_{L,R}$ must be defined by considering the image distance as *f'* not *f*. The denote *W*, in Eq. (6), has meaning of the size of magnified image on the screen. Eq. (7) and (8) show the formulas to calculate the image coordinates ($W'_{\mathrm{L,R}}$) and its disparity (DIS') on the screen when non-infinity of FOD is applied. In Eq. (7) to (8), note that the prime notation which located in each formula. The notation presents that the values are obtained by camera system having the case which FOD is non-infinity. If there has not the notation, we can use these equations to obtain image coordinates ($W_{\mathrm{L,R}}$) and their disparity (DIS) when FOD is infinity such as pin-hole system. In Fig. 5, the scheme shows two cases of reconstructed depth cues by the each disparity in the stereogram on the screen. One is that presenting the coordinates of depth cue which denoted as (X^I, Z^I) and the position of reconstructed depth image is presented as a dot on the screen. As the same way, the other is that presenting the coordinates of other depth cue which denoted as (X'^I, Z'^I) and the depth image is presented as a dot in front of the screen. In spite of the same object is used in the process, the former is caused when FOD is infinity and the later is caused when FOD is non-infinity.

Hence, the difference of both depth cues makes the contradiction to the result in the reported studies. The Eq. (9) to (10) shows that the formulas to calculate qualitatively the coordinates of reconstructed depth cue. As the same way in Eq. (7) and (8), the notation as a prime in Eq. (9) and (10) has charge of FOD status which is infinity or not.

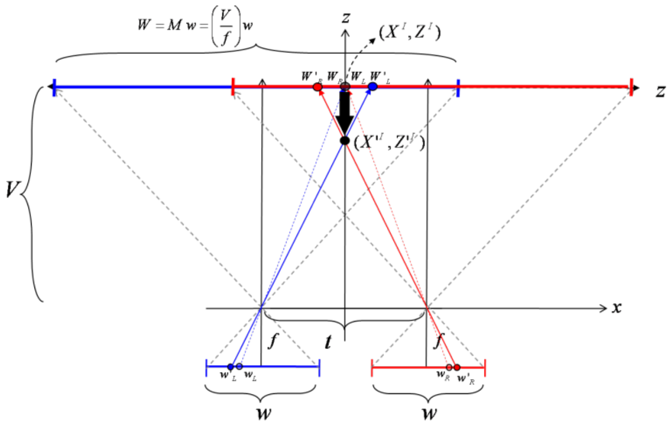

Fig. 5. The scheme shows two cases of reconstructed depth cues by the each disparity in the stereogram on the screen.

$$W'_L = M(-w'_L) - \frac{t}{2}\ ,\ W'_R = M(-w'_R) + \frac{t}{2} \quad \cdots\cdots\ (7)$$

$$DIS' \equiv (W'_L - W'_R) = M(dis') - t \quad \cdots\cdots\ (8)$$

$$Z'^I = \frac{V e}{(DIS' + e)} \quad \cdots\cdots\ (9)$$

$$X'^I = \frac{e(W'_L + W'_R)}{2(e + DIS')} \quad \cdots\cdots\ (10)$$

The graph, in Fig. 6, shows that the result of depth distortion when the practical data are substituted in Eq. (9) and (10). The practical data are constituted as following that focal length of lens (*f*) is 50mm, the width of detector in camera (*w*) is 38mm, the viewing distance (*V*) is 10m, the inter-distance (*t*, *e*) is 65mm, the magnification (*M*) is defined by *V/f* and object space is constructed in horizontal and vertical directions which having the each range from -5m to +5m horizontally and from 0m to 10m vertically to the origin of global coordinates. And the FOD is divided by four steps such as 0.5m, 1m, 5m and infinity to verify that its values make a depth distortion. Each type of line on the graph in Fig. 6 presents the result of depth cue when the stereogram is observed by an observer having viewing distance *V*. They are presented in ordered steps of FOD such as the line, the thin line, the two dots chain line and the bold dot line. As the result, the depth distortion becomes increasingly when FOD has shorter than infinity and the location of the object has so far in vertical direction. In addition, there is not occurred position distortion in horizontal direction because the ratio of increment of disparity on the stereo image due to variable FOD is constant. For this result, we can see that the only depth distortion in vertical direction to be generated due to the variable FOD.

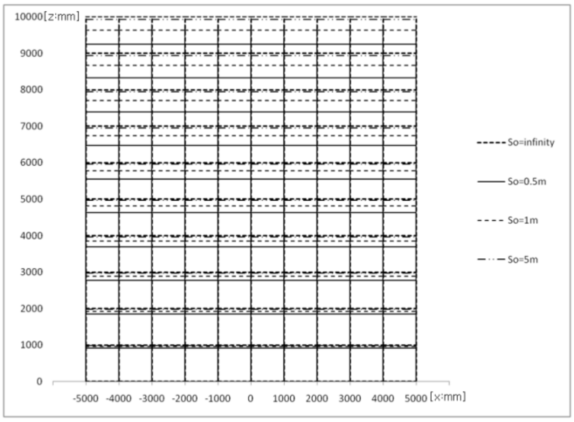

Fig. 6. The result of depth distortion when the practical data are substituted in Eq. (9) and (10).

3. Compensation of depth distortion

The cause of occurring depth distortion in stereoscopy based on camera system is turned out by the reasons that the FOD was not considered to obtain a stereo image for the object and the value of FOD makes it to increasing disparity in the stereo image. The increased disparity is strongly related to the ratio of image magnification. To remove the cause which occurring depth distortion, therefore, if we make it suitably to defining the ratio of image magnification then depth distortion will be compensated. In Fig. 7, the scheme of compensating depth distortion is drawn. To compensate depth distortion, the ratio of image magnification (M^C) which was defined as a function of FODs is suitably considered. Eq. (11) present the formula to calculate M^C and Eq. (12) and (13) present each the formula which are applied by M^C to calculate the compensated image coordinates on the screen ($W^C{}_{L,R}$) and their disparity(DIS^C). As we can see that the size of image and the disparity will be reduced because M^C is smaller than *M*. For this result, we can see one fact that the size of stereo image which was considered by reduced magnification for compensating depth distortion will be loss of a part as an edge. Is spite of the loss of size of image, this compensating depth distortion method is so simple and so powerful.

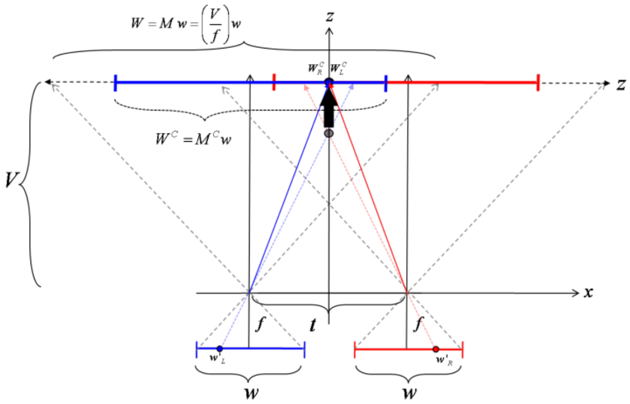

Fig. 7. The scheme shows the geometric analysis of the process for compensating depth distortion

$$M^C = M\left(\frac{s_o - f}{s_o}\right) \quad \cdots\cdots\ (11)$$

$$W_L^C = M^C(-w'_L) - \frac{t}{2}\ ,\ W_R^C = M^C(-w'_R) + \frac{t}{2} \cdots\ (12)$$

$$DIS^C \equiv \left(W_L^C - W_R^C\right) = M(dis) - t \ , \ \left\{ \text{if } \ f' < 0, \ dis' = \frac{f't}{z} \right\} \quad \cdots\cdots\cdots (13)$$

In Fig. 8, the graph shows that the depth distortion in displayed space is compensated and the graph is drawn by same way as the Fig. 6. For this graph, all of line is superposed on the line. The line presents the depth cue of the case when the stereo image was formed by FOD has infinity. Therefore, we can take the results that the depth distortion will be compensated by using the ratio of suitable magnification which was applied to the process and the value of compensated disparity in the stereo image will be equal to the disparity which is calculated by pin-hole model, because they will have had a invariant image distance as *f*. The Eq. (14) and (15) shows that the formulas to qualitative calculate the coordinates of compensated depth cue. As the same way in Eq. (9) and (10), the notation as a prime in Eq. (14) and (15) has charge of FOD status which is infinity or not. In addition, as the same case of Fig. 6, there is not occurred position distortion in horizontal direction too and we can see that the only depth distortion in vertical direction was compensated.

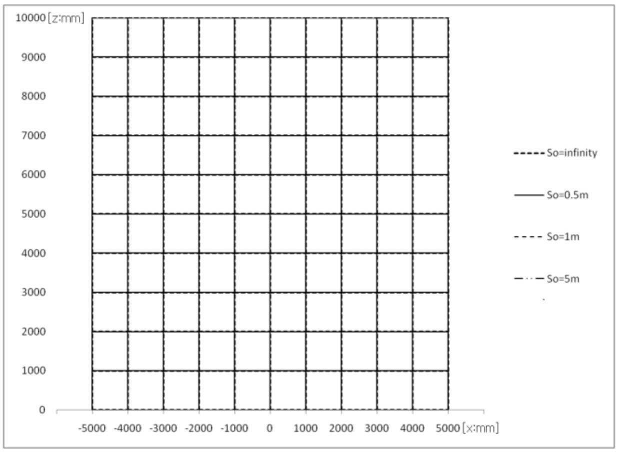

Fig. 8. The graph shows the result that the depth distortion is compensated by using the ratio of suitable magnification.

$$Z^{CI} = \frac{Ve}{\left(DIS^C + e\right)} \quad \cdots\cdots\cdots (14)$$

$$X^{CI} = \frac{e\left(W_L^C + W_R^C\right)}{2\left(e + DIS^C\right)} \quad \cdots\cdots\cdots (15)$$

4. Conclusion.

In this paper, correct depth analysis in stereoscopy is proposed. The cause of generating incorrect depth cue which mentioned as contradiction in the reported studies is analyzed by our investigation. The investigation is performed as following that the depth distortion is defined by using qualitative analysis on the viewpoint of geometry and the correct depth cue is obtained by using the formulas to compensate incorrect factors such as the ratio of magnification, the value of disparity and the variable image distances. These factors are strongly depends on the state of focused object distance. If we use camera system to perform the stereoscopy process, the state of FOD must be considered in. Since the value of FOD affect on the image distance in camera system and then, the image distance to be a cause which generating depth distortion. Consequently, the way compensating incorrect depth cue as depth distortion in reported studies is performed that applying the suitable ratio of image magnification which is considered the value of FOD to the process of stereoscopic analysis. Additionally, as we described that even if the stereo image had a correct depth cue by our method, there is a disadvantage that the image had loss of part as an edge but it a negligible factor. From what has been discussed above, we can conclude that our investigation is more adequate method to obtain correct depth cue in stereoscopy than the existing method. This investigation will change the direction of the entire study of stereoscopy analysis.

Reference

[1] Kwang-Hoon Lee, Soo-Ho Kim, Young-Soo Youn, Jung-Young Son, Min-Chul Park and Sung-Kyu Kim, “Constraints for stereoscopic vision in stereo camera phone”, SPIE Optics East ’07, V6778.

[2] Daniel B. Diner and Derek H. Fender, *Human Engineering in Stereoscopic Viewing Devices, Plenum press, New York and London,* p35-66

[3] Rudolf Kingslake, *LENS DESIGN FUNDAMENTAL - Chap 3, Academic press,* New York, p40-71

[4] Joseph M. Geary, *Introduction to LENS DESIGN - Chap 4, Willmann-Bell,* Virginia, p33-42

[5] Jung-Young Son, Kyung-Hoon Cha, Sung-kyu kim, Min-Chul Park and Sung-Hwan Jang, "Stereo-photography with handphone",SPIE ITCom’06, V6382-6.

[6] A. Woods, T. Docherty and R. Koch "Image Distortions in Stereoscopic Video Systems", Proc. SPIE V1915, pp36-48, 1993.

[7] Fundamental of 3-D Imaging Techniques(Japanese Edition), T. Izumi(Supervisor), NHK Science and Technology Lab., Ohmsa, Tokyo, 1995, ch. 3, pp55-60.