Disparity Based Quality Function for

Autostereoscopic Display Devices

Jung-Young Son, Vladimir V. Saveljev*, Min-Chul Park**, Dae-Sik Kim*** and Sung-Kyu Kim**

School of Computer and Communication Engineering, Kyungsan, Kyungbuk, Korea

*Whole Imaging Lab. Hanyang University, Seoul, Korea

**System Engineering Div., Korea Institute of Science and Technology, Seoul, Korea

***DM Div. Samsung Electronics, Suwon, Korea

Abstract

A two dimensional quality function which counts both the number of mixed view images and disparity between images is derived based on the one dimension quality function which counts the number of mixed view images in multiview 3 dimensional imaging systems. This function predicts the quality of images with reasonable accuracy. This is proved experimentally.

1. Introduction

The compositions of images projected to viewers' eyes at the viewing zones of multiview 3 dimensional(3-D) imaging systems using microoptics plates such as arrays lenticular, parallax barrier, microlens and gratings as the viewing zone forming optics, are different for different spatial regions in the zone. In general, the viewing zones are the spatial places where viewers can perceive depth sense through the images displayed on the display panel and they are geometrically divided into many spatial viewing regions. Each of these viewing regions is also divided into many sub-regions corresponding to the number of pixel cells in the display panel[1]. The compositions in the viewing regions are classified into the number of different view image composing the images. In the viewing regions along the parallel plane to the panel, at which the viewing distance is defined, individual view images are separately viewed but in the regions adjacent to the individual view image viewing regions, a part from each of two adjacent view images are patched together to form a new image, and in the viewing regions adjacent to these viewing regions, a part from each of 3 adjacent view images are patched. In this way, n different view images are patched together to compose a complete image at each of the viewing regions located further away from the parallel plane when n different view images are displayed at the panel. Hence the quality of projected images in these regions will be different for different regions. The image quality is more deteriorated as the number of different view images involved to form the new images increases[2]. In this point of view, a quality function was devised to quantify the image quality by simply taking the inverse of the number of different view images patched together[3,4]. However, the function did not consider the quality deteriorations due to increase in the widths of the viewing regions. The widths of viewing regions are either increased or decreased depending on the viewer distance from the parallel plane. They are increased if the distance

Three-Dimensional TV, Video, and Display VI, edited by Bahram Javidi, Fumio Okano, Jung-Young Son,
Proc. of SPIE Vol. 6778, 677804, (2007) · 0277-786X/07/$18 · doi: 10.1117/12.735751

increases but decreased if viewers are close to the display panel. If the widths are more than the viewers' interocular distances, the viewers' two eyes can be in the same viewing region. So viewers may not perceive any depth sense. Since the main quality of 3-D images is the depth sense, perceiving no depth sense is the most serious quality deterioration in 3-D images. For the case when the widths are decreasing, the viewers' two eyes are not located at two adjacent viewing regions but two viewing regions separated several viewing regions apart. In this case, the depth sense can be enhanced or deteriorated due to difficulty in fusing the images projected at the regions.

In this paper, a two dimensional quality function is derived to include the disparity between images on the display panel and its effectiveness is proved experimentally.

2. Image compositions at each viewing region

Since the image compositions at different viewing regions in the 3-D imaging systems are defined geometrically, the basic optical geometry of forming viewing zone in the 3-D systems should be examined to define the compositions. Fig. 1 depicts the viewing zone forming geometry of the 3-D imaging system for the case when the display panel is consisted of 5 pixel cells and each pixel cell is composed of four pixels specifying four different view images of 1, 2, 3 and 4. The number combinations in each viewing regions and viewing sub-regions represent a pixel of each view image in each pixel cell seen at each viewing sub-regions. Hence they represent the image composition seen at these regions. The parallel plane is the place where all magnified images of pixel cells on the display panel are completely superposed together and is named as viewing zone cross-section(VZCS). Since the pixel cells are having the same shape and pixel structure, and pixels in the panel are the same, all magnified pixels in each pixel cell are completely matched with those of other pixel cells in the panel. However since the magnified images are continuously expanding, the images are no longer matched at the viewing regions of before and behind the VZCS. Hence the compositions of images projected to viewer eyes will be different for different viewing regions. At the regions along the VZCS, there are four different viewing regions corresponding to four different view images. These viewing regions are specified as region 1. In these regions, the pixels belong to the same view images from all five pixel cells are seen. Hence different view images are separately seen at these regions. Behind and before the region 1, regions 2 for three different viewing regions, 3 for two different viewing regions and 4 for one different viewing region, are followed. The region 2 is also divided into sub-viewing regions corresponding to the number of pixel cell-1. Region 3 will be different for different number of the pixel cells. The image compositions of four sub-regions of the top viewing region in region 2, are either 11112, 11122, 11222 and 12222, or 21111, 22111, 22211 and 22221 from the top, respectively. For the second and third viewing regions, the image compositions at the viewing sub-regions are 22223(32222), 22233(33222), 22333(33322) and 23333(33332, and 33334(43333), 33344(44333), 33444(44433) and 34444(44443), respectively. Hence in this region, two adjacent view images are involved to form a new image at each viewing region. In region 3, the image compositions in sub-regions of top and bottom viewing regions are 11223(32211), 11233(33211), 12223(32221) and 12233(33221), and 22334(43322), 22344(44322), 23334(43331) and 23344(44332), respectively. In this region, 3 three adjacent view images are involved to form a new image for each viewing region. In region 4, the compositions are

11234(43211), 12234(43221), 12334(43321) and 12344(44321) for four sub-regions. In general, for n different view images with m ($m \geq n$) pixel resolution(pixel cells) for each view image, 1) there will be n^2 viewing regions and 2) these viewing regions are grouped into $2n-1$ regions, 3) regions No. 2 and n are divided into $m-1$ sub-viewing regions, and the projected image in each region is synthesized by the number of different view images corresponding to the region number. The geometry can be equally applied for IP[5]. Fig. 2 is showing the methods of synthesizing a new image by patching image strips from eight different view images. In this figure, it is assumed the strips have the same size but in actual case as shown in Fig. 1, the strip width is not the same. The width is different for different sub-regions. Fig. 3 shows examples of the patched image at the region 8 according to the procedure shown in Fig. 2. Fig. 3(a) is for forward order and 3(b) for reverse order. The reverse order images are seen at the viewing regions in left side of the VZCS. As shown in the figures, it is hard to find the difference of the new images from the individual view images and between them visually, though it is known that these images are synthesized from an image strip of each view images. This is why it is necessary to devise a new quality function which represents more effectively quantifying the quality of images projected to viewers' eyes in the viewing regions of the viewing zones.

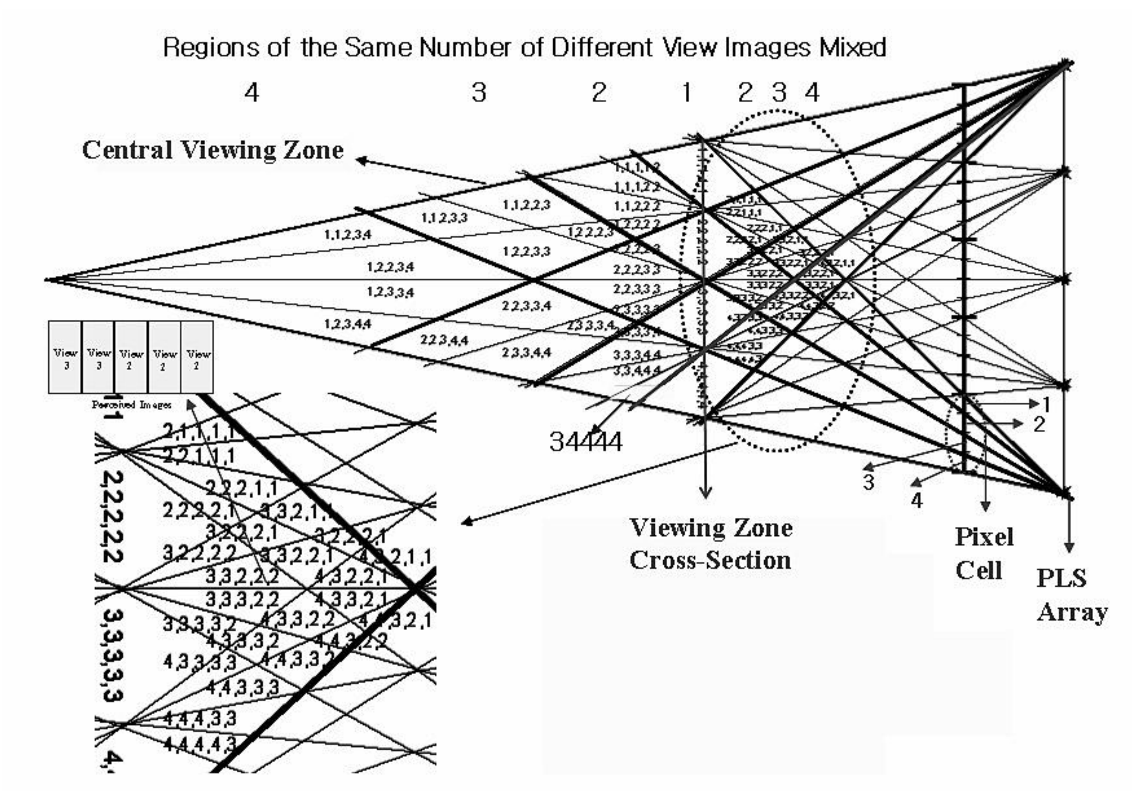


Fig. 1 Image compositions at different viewing regions in viewing zone. Regions 1, 2, 3 and 4 indicate numbers of different view images mixed together to form images seen at the regions..

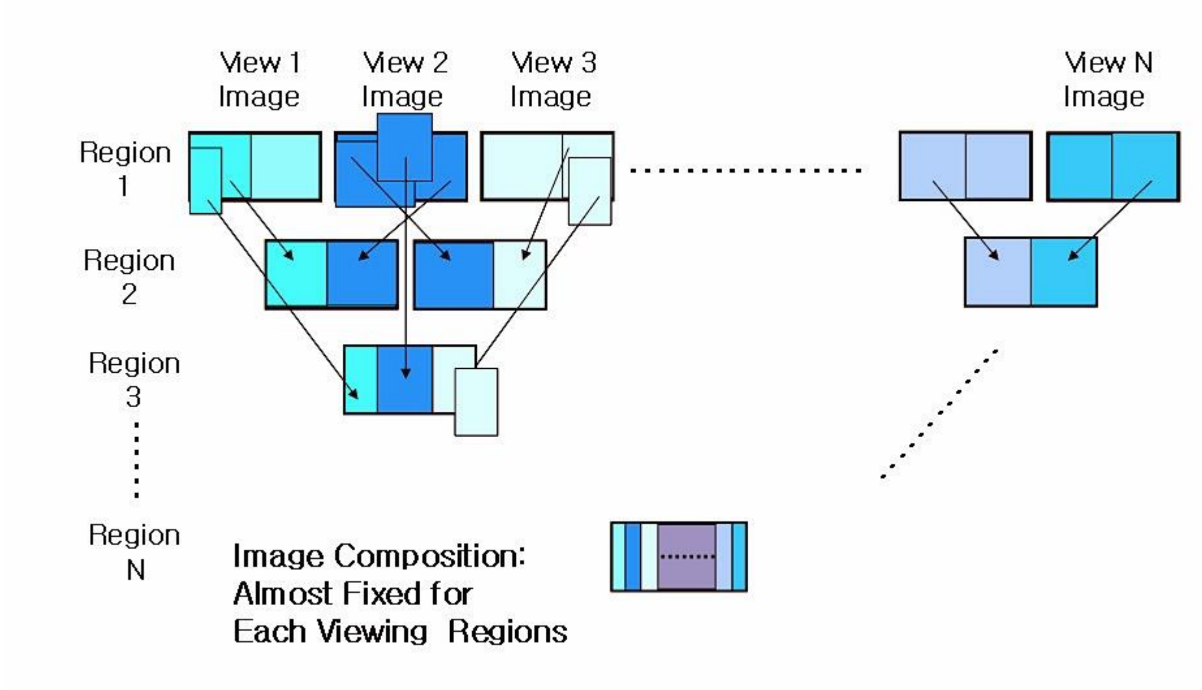
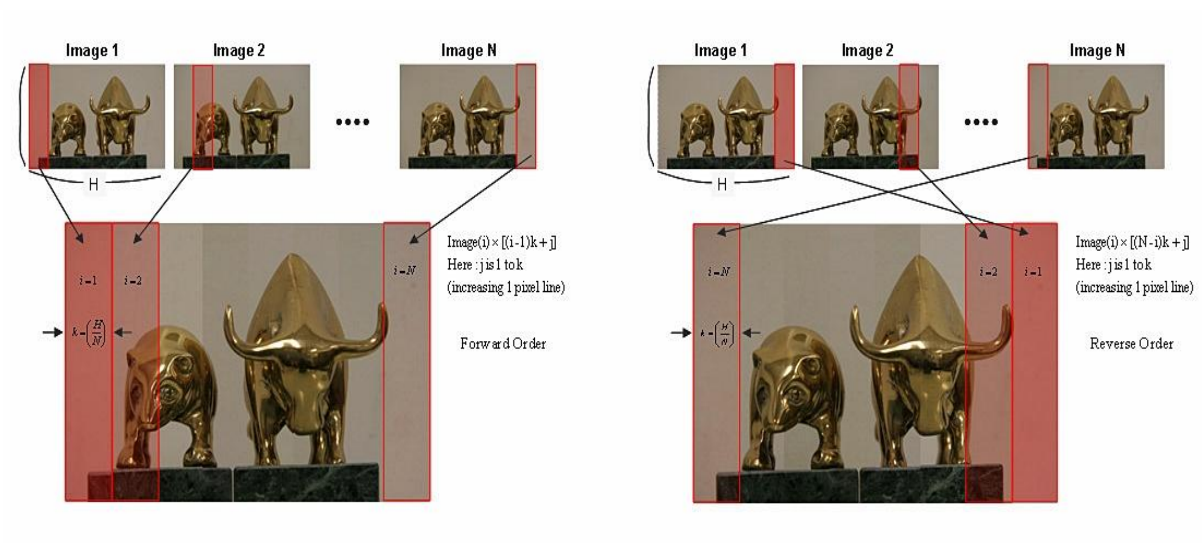


Fig. 2 Methods of synthesizing new images at different regions by patching image strips from different view images



(a) Forward Order (b) Reverse Order

Fig. 3 Examples of mixed images

3. A new quality function

The quality function $Q(j)$ (where $j=-n,...,0,...,n$) which is defined as the inverse of the number of different view images patched together, is represented as[3,4],

$$Q(j)=\frac{1}{1+|j|} \tag{1}$$

where $j=-n,...,0,...,n$ represents node numbers defined in Fig. 4. Fig. 4 is a redrawn figure of Fig. 1 without point light source array for $n=4$. It shows node planes and images seen at these node planes are specified by numbers. Numbers 1, 2, 3 and 4 represent different view images. If j is replaced by $i(i=j+n)$ as in Fig. 4, i becomes $i=0,...,2n$. In this case, the distance of each nodes from the display panel, $z_b(i)$ and the distance between nodes in each node, $w_b(i)$ are represented as,

$$z_b(i)=\frac{i}{n(c^{-1}+1)-i}\frac{1}{1+c}d$$

$$w_b(i)=\frac{c^{-1}}{n(c^{-1}+1)-i}b=\frac{1}{1+c-\frac{c}{n}i}\frac{b}{n} \tag{2}$$

where $c=b/D$, b and d are the width of VZCS and the distance between the display panel and VZCS, respectively. From Eq. 2, i is solved, transformed to j and is substituted to Eq.1. Then Eq. 1 becomes a function of z ($z_b(i)$ is replaced by z), i.e., $Q(j)$ becomes $Q(z)$.

$$Q(z)=\frac{z+\frac{1}{1+c}d}{\left(1-n\frac{1}{c}\right)z+(1+n)\frac{1}{1+c}d};\quad z<d \quad \text{and} \quad \frac{z+\frac{1}{1+c}d}{\left(1+n\frac{1}{c}\right)z+(1-n)\frac{1}{1+c}d};\quad z>d \tag{3}$$

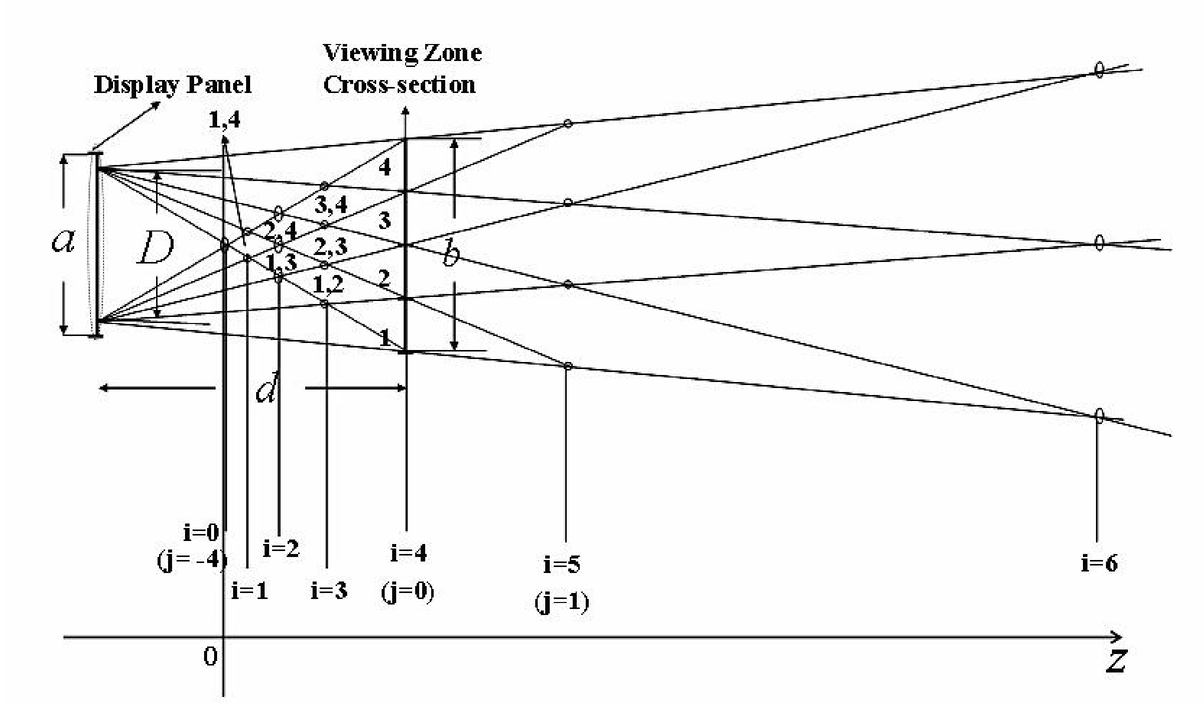


Fig. 4 The geometry defining disparity and node planes and points

Since $c = b/D$ and D is a variable representing the distance between two arbitrary horizontal pixel cells on the display panel, $Q(z)$ can be replaced by $Q(z, D)$. Hence Eq. 3 is written as,

$$Q(z,D) = \begin{cases} \dfrac{z + \dfrac{D}{D+b} d}{\left(1 - n\dfrac{D}{b}\right) z + (1+n)\dfrac{D}{D+b} d}, z < d \\ \dfrac{z + \dfrac{D}{D+b} d}{\left(1 + n\dfrac{D}{b}\right) z + (1-n)\dfrac{D}{D+b} d}, z > d \end{cases} \tag{4}$$

In Eq 4, D defines both the positions of node planes and a view image pixel from each of the two pixel cells, which will be combined at each node in each node plane, hence D is defined here as disparity.

In Fig. 5, Eq. 5 is depicted for 5 different D values of 0, 1, 2, 3, 4cm, $d = 25cm$, $b = 5.5cm$ and $n = 4$. The horizontal line represents z values in cm unit.

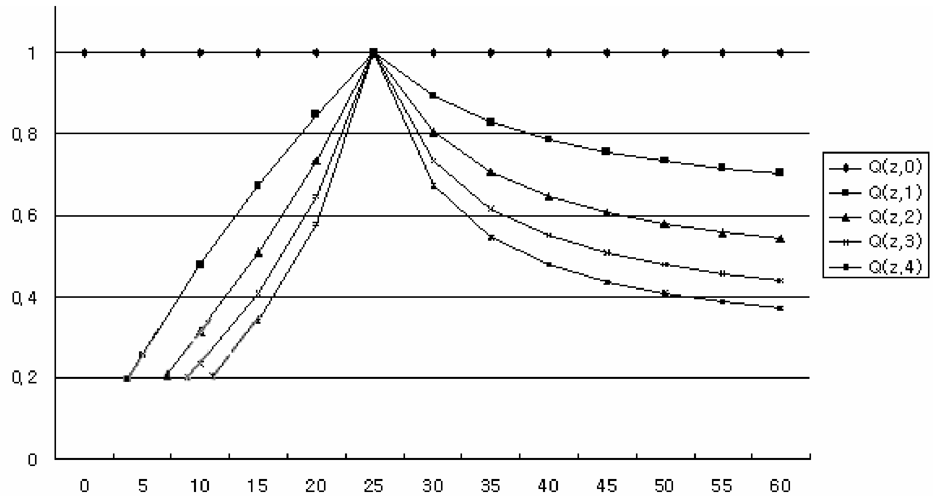


Fig. 5 $Q(z,D)$ for different D and z values

From Fig. 5, it is obvious that $Q(z,0)=1$ without regard to z values because only one pixel cell is involved in this case. The starting point of each $Q(z,D)$ is becomes different due to the fact that $i=0$ position becomes further away from the display panel as shown in Fig. 5. The $Q(z,D)$ values for the starting points of all D values are the same as $1/(n+1)$ as expected. Since the z value ranges for higher D values are decreasing, the area for higher $Q(z,D)$ values is decreasing. This means that the images with big disparities can only be observed with a good quality, i.e., with no or small distortions within viewing regions in a short range of depth direction. Out of the range, they will be seen with a low quality, with higher distortions. As a consequence, some visual images with exceeding disparity can be discarded from practical usage even if they could be built with using a formal procedure. The upper limit for disparity can be found from Eq. 5. But still human factor-wise research is necessary to determine the limit.

4. Experimental Results

The experiments are performed with a portable 3-D display device[6] capable of displaying 4 different view images in the horizontal line. Virtually the device contains 120 light sources. Totally, 9 experiments were made for the different combinations disparity values and nodal plane positions. A testing image consisted of four vertical lines is located near the center of the screen. Each vertical line represents a view image and each view image is identified by the relative position in the screen. Since the lines are thin, for better identification of each line, the testing image is enclosed within a bounding square. This square works like a pixel cell. The vertical lines are displayed independently so that they do not interfere with each other. To verify the image mixing at different node planes for different disparity values as shown in Fig. 4, another square with the same vertical lines and size as in the first square is displayed with a certain distance D in the horizontal direction. This testing layout allows evaluating experimentally the composition of images viewed at the different viewing regions and contribution of each pixel cell in different locations of the panel for the device, to the composition. The experiments were performed in the following way: A camera movable to both horizontal($x-$axis) and depth($z-$axis) directions is installed in front of the device to make its axis normal to the display panel of the device. Then the two squares with controlled distances are displayed and seen by the camera. The camera position is adjusted to both directions for better images. In Table 1, the theoretically obtained x and z axes values from Eq. 4 are compared with the experimentally obtained values. As shown in Table 1, the difference between the experimental and the theoretical are less then 8%. This means that Eq. 4 can predict the image quality fairly well.

Table 1. Comparisons between theoretically and experimentally obtained x and z values

Disparity *D*, cm	**Combinations**	**Theoretical**		**Experimental**	
		Distance *z*, cm	Displacement *x*, cm	*z*, cm	*x*, cm
1.0	{1,2,3,4}	25.0	1.38	24.0	1.3
1.0	{1+2, 2+3, 3+4 }	10.5	0.58	13.5	0.75
2.0	{1,2,3,4}	25.0	1.38	26.0	1.25
2.0	{1+2, 2+3, 3+4, 1+4}	14.8	0.94	17.0	0.9
3.0	{1,2,3,4}	25.0	1.38	28.0	1.4
3.0	{1+4}	17.1	0.94	20.5	0.9
3.0	{1+3, 2+4}	13.0	0.72	14.5	0.8
3.0	{1+2, 2+3, 3+4, 1+4}	10.5	0.58	12.4	0.6
4.0	{1,2,3,4}	25.0	1.38	27.0	1.45
Average		25.0	1.38	26.3	1.33
Standard deviation				1.71 (6.5%)	0.96 (7.2%)

The last two rows of Table 1 contain the values averaged for the entire experiment. In the current experiments, the averaging is only applied to the combination {1,2,3,4} observed from the design distance.

Fig. 6 shows the clear appearance of 4 view images (combination {1,2,3,4}) from 4 corresponding locations within the designed observation base, i.e., at VZCS. In this case, only a square is displayed. There is no disparity.

Fig. 6. Different view images seen at VZCS .

In Fig. 7 and 8, various combinations of views are shown as they are seen from locations other than the design base. The photographs were taken from transition areas of different types, mixed with neighbor (combinations 1+2, 2+3, etc.) and mixed with neighbor of neighbor (combinations 1+3, 2+4, etc.).

Fig. 7. Images seen at various viewing positions when $D = 2cm$ and $z = 17cm$

Fig. 8. Images seen at various viewing positions when $D = 3cm$ and $z = 14.5cm$

These figures clearly demonstrates that the predictions of different view image mixing as depicted in Fig. 4 are right.

Conclusion

The quality function represented by Eq. 4 can predict the image compositions at different viewing regions with a good accuracy. In current experiment, the prediction is correct with less than 8% difference from the experimentally obtained values. The function predicts that as the disparity increases the viewing zone depth for a specified quality value decreases. The effectiveness of the quality function should be checked by considering human factors

References

1. Jung-Young Son, Vladmir V. Saveljev, Yong-Jin Choi, Ji-Eun Bahn and Hyun-Hee Choi, “Parameters for Designing Autostereoscopic Imaging Systems Based on Lenticular, Parallax Barrier and IP Plates”, *Optical Engineering, V42, No.11, pp3326-3333, (2003)*

2. Jung-Young Son, Vladmir V. Saveljev, Jae-Soon Kim, Sung-Sik-Kim and Bahram Javidi, **"**Viewing Zones in 3-D Imaging Systems Based on Lenticular, Parallax Barrier and Microlens Array Plates**",** Appl. Opt. 43(26), 4985-4992(2004): Appl. Opt. Cover

3. J.-Y. Son, V. V. Saveljev, Y.- J. Choi, S.-K. Kim, “Quality Quantification of Multiview Three Dimensional Image”, IDW’02, Proceedings of The Ninth International Display Workshop, Hiroshima, Japan, pp1225-1228, 2002

4. V. V. Saveljev, Jung-Young Son and Kyung-Hoon Cha, “Estimation of Image Quality in Autostereoscopic Display,” SPIE V5908, pp590807-1(-14), 2005(2005, Annual Meeting)

5. Jung-Young Son, Shin-Hwan Kim, Dae-Sik Kim and Bahram Javidi, “Image Forming Principle of Integral Photography,” will be published in IEEE/OSA Journal of Display