

Digital Magnification of 3D Imaging Using an Intermediate-View Reconstruction Based on an Adaptive Block Matching Algorithm

Kyung-Hoon Bae

Communication/Platform CFCS Group, Samsung Thales Co., Ltd., San 14-1, Giheung-gu, Yongin,
Gyeonggi 446-712, Korea
E-mail: khbae.bae@samsung.com

Dong-Choon Hwang

Samsung Electronics Co., Ltd., Yongtong-gu, Suwon, Gyeonggi, 443-370, Korea

Eun-So Kim

3D Display Research Center, Department of Electronic Engineering, Kwangwoon University,
447-1 Wolge-Dong, Nowon-Gu, Seoul 139-701, Korea

Abstract. In this paper, digital magnification of three-dimensional (3D) imaging, using an ABMA intermediate view reconstruction (IVR) based on an adaptive block matching algorithm is proposed. The number of the elemental images obtained from a one-step pickup process can be computationally increased through the use of the proposed IVR without the need for mechanical movement and a long multistep pickup process. To show the feasibility of the proposed magnified 3D imaging system, using test images of a car, bottle, phone, and watch, it was shown that the proposed ABMA algorithm improves the peak signal to noise ratio of reconstructed intermediate images by as much as 5.16 dB on average in comparison with a conventional block and feature-based algorithm. © 2008 Society for Imaging Science and Technology.
[DOI: 10.2352/J.ImagingSci.Technol.(2008)52:5(050504)]

INTRODUCTION

Integral imaging (II), as originally proposed by Lippmann in 1908, has been researched actively as a promising technique for next generation three-dimensional (3D) imaging and displays.^{1,2} In II, 3D objects with full parallax and continuous viewing points can be recorded and reconstructed. Another method involves the use of a moving array-lenslet technique (MALT)³ to pickup many time-multiplexed elemental images instead of a modification of elemental images. Jang and Javidi reported that scaling can be accomplished by controlling the spatial ray sampling rate in the pickup process of II. The use of MALT can provide the uniform scaling of 3D objects between lateral and longitudinal spatial coordinates. The display of 3D images is performed with a stationary lenslet array. However, this requires a multistep pickup process via the vibration of a lenslet array in the pickup part. Therefore, mechanical movement and a

long pickup time make it difficult to implement a scalable real-time II system.

In this paper, a new scheme for magnified 3D images⁴ without mechanical movement or a long multistep pickup process in a II system is proposed. In the proposed scheme, the number of the elemental images obtained from a one-step pickup process can be computationally increased using intermediate view reconstruction (IVR).⁵⁻⁷ The IVR technique is now actively studied as a method to implement practical true-view stereoscopic display systems without an increase in the number of cameras used or image data to be processed and transmitted. In the process of intermediate view reconstruction, many matching algorithms (such as block, pixel, and feature-based) are presently used for disparity estimation. Generally, a block-based matching algorithm (BMA) is known to have a fast execution time, but not an accurate matching ability. The feature-based matching algorithm (FMA) has the capability of accurate matching at the edges of an object, but its overall execution time depends greatly on the characteristics of the input image. In these conventional disparity estimations, the matching window size is fixed as a pixel or a block for given stereo input images. For this reason, their processing time is excessively long. However, this technique can be used more effectively in practical applications if the matching window size is adaptively selected according to the local characteristics of the input stereo image pair. In this approach, pixel and block-based matching algorithms are selectively and complementarily taken in each local area depending on the specific characteristics of the input stereo image, which improves the overall performance of the disparity matching. Here, image complexity is characterized as the local gray levels of the corresponding matching points between the stereo image pairs. Additionally, these local gray levels are assigned as feature values indicating image complexity at those points.

Received Jan. 22, 2008; accepted for publication Jun. 6, 2008; published online Sep. 23, 2008.

1062-3701/2008/52(5)/050504/4/\$20.00.

Therefore, this paper proposes an adaptive block matching algorithm (ABMA) for effective IVR. In the ABMA, the pixel, block, and feature-based matching algorithms are selectively and complementarily taken in each local area depending on the specific characteristics of the input image. This has the effect of improving the overall performance of the disparity estimation. The proposed IVR makes it possible to synthesize the required number of intermediate elemental images digitally using only optically elemental images of a 3D object that had undergone a pickup process. Essentially, elemental images from a one-step pickup cannot sufficiently magnify 3D reconstructed images in a conventional II system; hence, the proposed IVR is introduced here to increase the number of elemental images. This has the same effect as an increase in the spatial ray sampling rate when MALT is used.

Accordingly, the proposed system can be used as a solution for a real-time scalable II system without the need for additional time for the pickup of a 3D object and mechanical movement of a lenslet array. This is the first known use of IVR in an optical II application.

MAGNIFIED 3D IMAGES BY ABMA-BASED IVR

Disparity Estimation

The BMA is used as a disparity estimation method in which, once one of the stereo image pairs is divided by a block unit, each block in the image (left image) undergoes a matching process with the individual blocks of the other image (right image) by means of a cost function.

Equation (1) shows the MAD function used for disparity estimation in this paper. In the equation, N_x, N_y denote the size of the block, and $I_L(i, j)$ and $I_R(i + k, j + l)$ indicate the coordinates of the left image and the corresponding coordinates of the right image, where k and l are horizontal and vertical disparity, respectively,

$$\text{MAD}(i, j) = \frac{1}{N_x N_y} \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} |I_L(i, j) - I_R(i + k, j + l)|. \quad (1)$$

First, the left image is divided by the block with $N_x \times N_y$ pixels. For each block, the cost functions of Eq. (1) are then calculated individually with the corresponding block of the right image, as the block is shifting within a given search range. Finally, a block that minimizes the cost function is determined as a corresponding matching block of the right image to the selected block of the left image, and the position difference between the blocks in the right and left images is assigned as a disparity vector.

Adaptive Block Matching Block Algorithm-Based Disparity Estimation

Figure 1 shows an algorithm of the ABMA. It adaptively determines the adaptive block size by the feature value⁶ that is compared with a predetermined threshold value. The feature values of the input stereo image that determine the proper matching window size can be extracted from the input stereo image pair. That is, by applying an edge detection algorithm to the input stereo image pair, the edge informa-

tion of this stereo image can be extracted. The extracted data is discontinuous and is displayed according to the value of the pixel brightness, depending on the characteristics of the input stereo image pair.

In the ABMA, feature values that are extracted from an input stereo image pair are normalized and the threshold value is optionally set as several steps in the range of 0–1. Correspondingly, an equal number of adaptive blocks with different sizes can be selected. As shown in Eq. (2), six steps of the threshold value (0.0, 0.2, 0.4, 0.6, 0.8, 1.0) and six adaptive blocks sizes (32×32 , 16×16 , 8×8 , 4×4 , 2×2 , 1×1) were selected for this study, where n , and Th_n represent an integer in the range of ($1 \leq n \leq 6$), and the threshold value of the n th step, respectively, as

$$\frac{1}{N_x N_y} \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} |I_L(i, j) - I_R(i + k, j + l)| > \text{threshold value}(\text{step}[n]) \Rightarrow 2^{6-n} \times 2^{6-n}. \quad (2)$$

When the feature values become smaller, a larger adaptive block size is used. The opposite is true when the feature values become larger.

Regularized Disparity Estimation

Several regions in each image have similar disparity vectors; however, in the matching process, though vertical vector similarity exists, only the similarity in the horizontal direction is considered. Accordingly, considering the vertical vector similarity, it is possible to remove the false disparity vectors. Therefore, in this paper, a disparity estimation algorithm employing a regularization scheme based on neighborhood averaging is used to alleviate the problems of matching window overlapping and misallocation that occur with the conventional disparity estimation. Here, the regularized pixel $g(x, y)$ is defined by Eq. (3)

$$g(x, y) = \frac{1}{M} \sum_{(i, j) \in s} f(i, j). \quad (3)$$

In this equation, M and S are the number and set of neighborhood pixels ($i \times j$), respectively. In this paper, Eq. (4) is used to preserve edge value in an edge region

$$g(x, y) = \begin{cases} \frac{1}{M} \sum_{(i, j) \in s} f(i, j); & \text{If } \left| f(x, y) - \frac{1}{M} \sum_{(i, j) \in s} f(i, j) \right| < T. \\ f(x, y); & \text{otherwise} \end{cases} \quad (4)$$

If the difference between the regularized and original pixel value $f(x, y)$ is smaller than a predetermined threshold, the original pixel value is replaced with the regularized value. In addition, during a right image reconstruction process, some occluded regions may exist in which one of the stereo cameras visualizes while the other does not. As a disparity vector is not allocated to this occluded region, the average value of the disparities of nearby regions is assigned through the process of disparity stability. If the viewpoint is not oc-

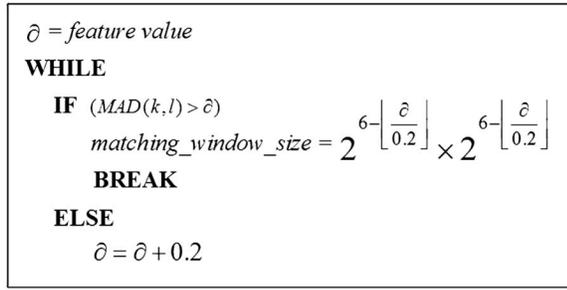


Figure 1. Flowchart of adaptive block matching algorithm.

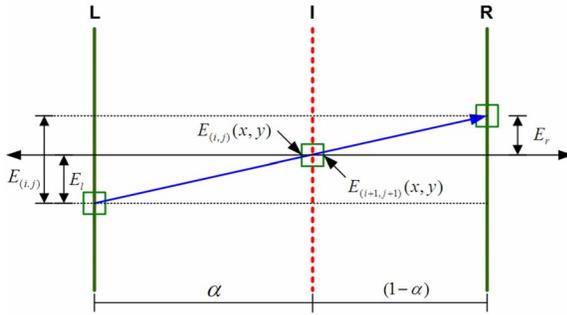


Figure 2. Concept of IVR.

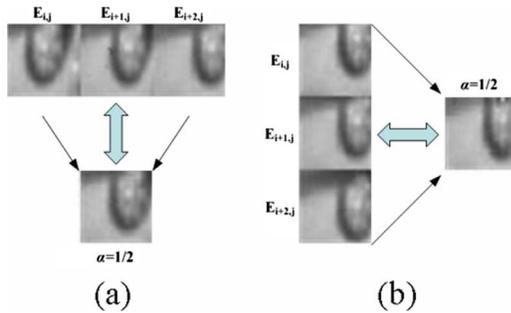


Figure 3. IEI calculated from the elemental images.

cluded at this stage, the disparity is defined as the distance between the image points in both images. Equation (5) shows the disparity in the horizontal direction and the relationship between the right image I_r and the left image I_l

$$I_r = \begin{bmatrix} i_r \\ j_r \end{bmatrix} = \begin{bmatrix} i_l + DV(i_b, j_l) \\ j_l \end{bmatrix} = I_l + \begin{bmatrix} DV(i_b, j_l) \\ 0 \end{bmatrix}. \quad (5)$$

Intermediate View Reconstruction

The proposed intermediate view images are synthesized using an interpolation method with a weighted mean value of the left and right images,⁷ as shown in Figure 2. The disparity vectors of these regions are then replaced with the mean values of the disparity vectors of the nearby regions by the process of disparity regularization. Through this full process of the proposed IVR algorithm, a more natural multiview and stereoscopic 3D image can be synthesized.

The distance from the left to the right image plane is 1, such that $\alpha \in [0, 1]$. For example, $\alpha=0$ represents the left image and $\alpha=1$ represents the right image, and the interval values represent the intermediate view image

Table I. Comparison to calculate a PSNR.

Method	PSNR[dB] of IEI			
	Car	Bottle	Phone	Watch
BMA	30.18	27.15	25.35	24.31
FMA	34.16	28.34	27.32	26.54
Proposed IVR	36.87	34.12	32.14	29.19

$$\hat{d}_{ij} = d_L + d_R = \alpha \hat{d}_{ij} + (1 - \alpha) \hat{d}_{ij}, \quad (6)$$

$$I_I(i, j) = I_L(i + \alpha \hat{d}_{ij}, j) = I_R(i - (1 - \alpha) \hat{d}_{ij}, j), \quad (7)$$

where \hat{d}_{ij} is the disparity value in the search range, I_I is the intersection at which the disparity of the left image and the right image cross, I_L is the block of the left-side image, and I_R is the block of right-side image corresponding to I_L . As shown in Fig. 2, due to the relationship expressed by Eq. (6), usually point $I_I(i, j)$ of the intermediate view image plane can be reconstructed from disparity vector \hat{d}_{ij} of the left, right, and distance α with Eq. (7). Moreover, in this paper, intermediate views that are more natural can be reconstructed using an interpolation scheme with a weighted average. Equation (8) shows the case of interpolation with a weighted average by position α of the viewpoint

$$I_I(i, j) = (1 - \alpha) \cdot I_r(i - \hat{d}_{ij}(i, j), j) + \alpha \cdot I_l(i - \hat{d}_{ij}(i, j), j). \quad (8)$$

Magnification Using MALT and IVR

The interlaced elemental images picked up from MALT are used for the magnification of the 3D integrated images. However, mechanical movement and a long pickup time prevent the implementation of a real-time system. To overcome this problem, an IVR is employed to increase the number of elemental images in a manner similar to the function of MALT. This IVR is computationally implemented with a computer, implying that no mechanical movement of the lenslet array in the pickup part is necessary and that a long pickup time does not occur.

The 3D object is picked up with the one-step process and the picked up elemental images are then transmitted to the display system. The display system is composed of an image processing part and a projection II system part. The former is used to generate intermediate elemental images (IEI) using ABMA-based IVR for magnification of the 3D images, and the latter is used to display 3D images to a lenslet array screen. The elemental images from the one-step pickup can be transmitted in real time because the proposed system is composed of commercial pickup devices of the type. In fact, the elemental images from the one-step pickup are not sufficient for use in the magnification process of the II system. Therefore, ABMA-based IVR is used to increase the number of elemental images.

It is important to note that the generation of intermediate elemental images takes place using IVR. It is assumed

that the (i, j) th elemental image is represented by $E_{i,j}(x, y)$, where x and y are the pixel positions within the elemental image. Here $i=1, 2, \dots, M$ and $j=1, 2, \dots, N$. If IVR is applied to two neighboring elemental images, $E_{i,j}(x, y)$ and $E_{i+1,j}(x, y)$, it is possible to obtain the intermediate views for various α values from Eq. (8). These are termed IEI. In the II system in this system, α is used as a magnification control parameter. If 3D objects are magnified n times, α then has $n-1$ different values with an interval of $\Delta\alpha=1/n$. Figure 4 shows an example of calculating IEI for three different α values with $\Delta\alpha=1/4$. To obtain the entire range of IEI, it is necessary to perform ABMA-based IVR for all elemental images in both the horizontal and vertical directions.

SIMULATIONS AND RESULTS

The elemental images were captured by a CCD camera through a pickup lenslet array. The object was positioned 3 cm from the lenslet array. The lenslet array uses 50×50 lenslets, and each lenslet is mapped with 30×30 pixels in the CCD camera. The focal length and diameter of the lenslet are 3 mm and 1.08 mm, respectively. To analyze the image quality of the calculated IEI, some comparisons were performed, as shown in Figure 3. Fig. 3(a) shows a comparison between neighboring elemental images horizontally. The reference image is required to compare the quality of the IEI. Therefore, three neighboring elemental images are considered: $E_{i,j}(x, y)$, $E_{i+1,j}(x, y)$, and $E_{i+2,j}(x, y)$. The IEI is calculated only between $E_{i,j}(x, y)$ and $E_{i+2,j}(x, y)$, and $E_{i+1,j}(x, y)$ is used as the reference image. In effect, $E_{i+1,j}(x, y)$ can be considered as the IEI optically obtained using the lenslet array. The calculated IEI was compared with the reference image. The average peak signal-to-noise ratio (PSNR) was found to be 33.08 dB after repeating the procedure for all of the i and j values. This shows that the calculated IEI can be used in the II system without significant degradation in the reconstructed 3D images. In contrast, a vertical comparison between neighboring elemental images is shown in Fig. 3(b).

The picked up elemental images are shown in Figure 4(a). They consist of 990×750 pixels. The enlarged elemental images for the tire of "car" are shown in Fig. 4(b). Here, it is known that each elemental image has its own perspective of a 3D object. IVR was applied to these elemental images. Fig. 4 shows the IEI calculated from the elemental images of Fig. 3 using three different α values of ($n=4$): $\frac{1}{4}$, $\frac{2}{4}$, and $\frac{3}{4}$.

Table I shows the PSNR of the intermediate view synthesized by each algorithm for the images. It was found that the proposed algorithm improves the PSNR by 6.69, 6.97, 6.79, and 4.88 dB over that of the BMA. Moreover, it also improves the PSNR by 2.71, 5.78, 4.82, and 2.65 dB over

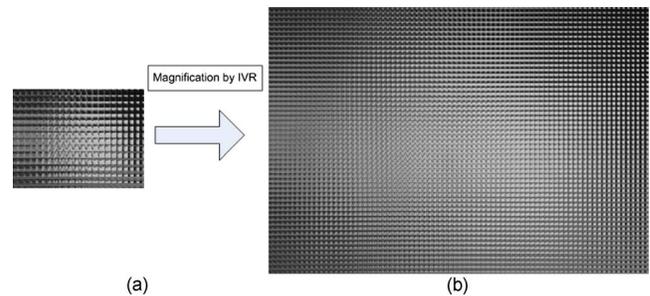


Figure 4. Experimental results for magnification: (a) the picked up EI; (b) the magnified elemental images times by ABMA-based IVR.

that of the FMA for each of car, bottle, phone, and watch images, respectively.

CONCLUSION

In this paper, digital magnification of 3D imaging by employing an ABMA-based IVR is proposed and experimentally demonstrated. From a number of successful experiments that centered on displaying scalable 3D reconstructed images for real 3D objects, the feasibility of the proposed II system is shown. From the experimental results, it was found that the proposed algorithm improves the PSNR by 6.69, 6.97, 6.79, and 4.88 dB over that of the BMA. It was also found that it improves the PSNR by 2.71, 5.78, 4.82, and 2.65 dB over that of the FMA for the images of the car, bottle, phone, and watch, respectively.

ACKNOWLEDGMENTS

This research was supported by MIKE (Ministry of Knowledge Economy), Korea, under the ITRC (Information Technology Research Center) support program supervised by the IITA (Institute of Information Technology Assessment) (No. IITA-2008-C1090-0801-0018).

REFERENCES

- G. Lippmann, "La photographie integrale", *Compt. Rend.* **146**, 446–451 (1908).
- F. Okano, H. Hoshino, J. Arai, and I. Yuyama, "Real-time pickup method for a three-dimensional image based on integral photography", *Appl. Opt.* **36**, 1598–1603 (1997).
- J.-S. Jang and B. Javidi, "Improvement of viewing angle in integral imaging by use of moving lenslet arrays with low fill factor", *Appl. Opt.* **42**, 1996–2002 (2003).
- R. Ponce-Diaz, R. Martínez-Cuenca, M. Martínez-Corral, B. Javidi, and Y. W. Song, "Digital magnification of three-dimensional integral images", *J. Disp. Technol.* **2**, 284–291 (2006).
- J.-W. Bae, H.-C. Park, E.-S. Kim, and J.-S. Yoo, "Efficient disparity estimation algorithm based on spatial correlation", *Opt. Eng. (Bellingham)* **42**, 176–181 (2003).
- K. H. Bae, J. J. Kim, and E.-S. Kim, "New disparity estimation scheme based on adaptive matching windows for intermediate view reconstruction", *Opt. Eng. (Bellingham)* **42**, 1778–1786 (2003).
- A. Redert, E. Hendriks, and J. Biemond, "Synthesis of multiviewpoint images at non-intermediate positions", *ICASSP-97* **4**, 2749–2752 (1997).