Improved Depth Perception of Single-view Images

Jae-Il Jung[1], **Jong-Ho Lee**[2], **In-Yong Shin**[3],
Ji-Hee Moon[4], and **Yo-Sung Ho**[5], Non-members

ABSTRACT

In many situations, we perceive 3-D depth feeling from 2-D images. Among various depth cues, monocular depth cues make us perceive depth from single-view images. In this paper, we propose several methods to enhance monocular depth cues and improve depth perception of single-view images. The proposed methods deal with four depth cues: linear perspective, aerial perspective, focus, and shadow effects. Once we have a 2-D colour image and its corresponding depth information, we can modify its depth cues and enhance 3-D depth perception in the 2-D image. Experimental results show that the proposed methods can enhance depth perception of single-view images without any considerable visual distortions.

Keywords: Monocular Depth Cue, Depth Perception, Single-View Image, Depth Enhancement

1. INTRODUCTION

With the rapid developments of technologies related to image processing, sensing and display, the interest in three-dimensions (3-D) broadcasting is rising. The main goal of 3-D broadcasting is to capture a 3-D scene and display it in a different place. In order to provide realistic feeling, 3-D broadcasting gives viewers various cues for depth perception.

Depth perception is the visual ability to perceive the world in 3-Ds. In the human visual system, the 3-D world projects onto the curved surface, the retina, at the back of the eye, but the retina is a two-dimensional (2-D) surface. Since we cannot have direct access to the depth of visual space, we estimate depth information based on two cues: binocular and monocular depth cues [1]. The binocular depth cues provide depth information when viewing a scene with both eyes and include stereopsis and convergence. Those are based on the simple fact that our eyes are located at different positions, and provide us extreme 3-D feelings.

In contrast, the monocular depth cues help us to perceive depth when viewing a scene with one eye, and contain linear perspective, occlusion, aerial perspective, and so on. Because we can estimate relative depth on the basis of monocular depth cues from a single-view image, the monocular depth cue is often called the pictorial cue.

Among both depth cues, many researchers have focused on the binocular depth cue to generate 3-D contents and to provide realistic images to viewers due to its extreme effects. Various types of approaches have been developed [2][3]. In order to appreciate 3-D contents based on the binocular depth cues, we should be equipped with special devices such as stereo and multi-view display devices, but the number of these special devices in use is not enough yet. Hence the methods based on binocular depth cues are not suitable for the present states. The alternative approach for 3-D broadcasting is enhancing monocular depth cues. The monocular depth cue cannot provide extreme effects as the binocular depth cue does, but it can be displayed on conventional display devices.

In this paper, we propose methods to enhance monocular depth cues. Among monocular depth cues, we especially focus on four main cues: aerial perspective, linear perspective, shadow, and focus effects. The proposed algorithm requires a color image and its corresponding depth map as input images. According to the distance, each monocular depth cue is modified and enhanced. The proposed algorithm is effective for conventional display devices, especially low-resolution display devices such as mobile phones and portable media players.

2. DEPTH PERCEPTION WITH CUES

We always sense the distance between objects, and this process is called depth perception [4]. It arises from a variety of depth cues, and the cues play an important role in the field of 3-D broadcasting. The cues are typically classified into two types according to the number of required eyes. While binocular depth cues require input from both eyes, monocular depth cues that require input from one eye.

2.1 Binocular Depth Cue

Binocular cues enable us to sense depth when watching objects with both eyes. There are two binocular depth cues: stereopsis and convergence.

Since we have two eyes placed at different position,

Manuscript received on , ; revised on , .

[1,5] The authors are with The Gwangju Institute of Science and Technology (GIST) 261 Cheomdan-gwagiro, Buk-gu, Gwangju 500-712, Korea Telephone: +82-62-970-2258 Fax: +82-62-970-3164, E-mail: jijung@gist.ac.kr and hoyo@gist.ac.kr

we can refer to information derived from the different projection of objects onto each eye. Based on these two images of the same scene captured at slightly different position, it is possible to triangulate the distance to an object. When an object is far away, the disparity of two images projected on both eyes will be small, and vice versa. This principal is known as stereopsis and is the main reason why we can perceive depth when viewing Magic Eyes, 3D movies and stereoscopic photos.

Another binocular depth cue is convergence come from the oculomotor behavior of two eyes. When the two eyes focus on the same object, they converge. The convergence will stretch the extraocular muscles, and the effect from this muscle helps depth perception. The angle of convergence is smaller when the eyes focus on far objects as can be seen in Fig. 1.

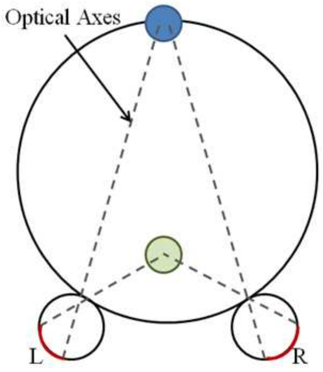


Fig.1: *Principal of Binocular Depth Cues*

Although we can experience the extreme 3-D feeling through the binocular depth cues, these cues are only enabled on special display devices.

2.2 Monocular Depth Cue

Dissimilar to binocular depth cues, monocular depth cues are available when viewing a scene with one eye. From them, we can perceive the depth of the object with only one eye even though the perception is worse than that with two eyes. Monocular depth cues include motion parallax, relative size, linear perspective, aerial perspective, occlusion, focus, shadow, and so on.

Linear Perspective. Linear perspective is the property that parallel lines converging at infinity allows us to reconstruct the relative distance of objects or landscape features. It can be explained as a phenomenon of Emmert's law[5] reported in 1881. A retinal image appears to change in size, according to whether it is seen as lying nearby or far away. This makes intuitive sense: an object of constant size will project progressively smaller retinal images as its distance from the observer increases. Figure 2 shows the principal of linear perspective.

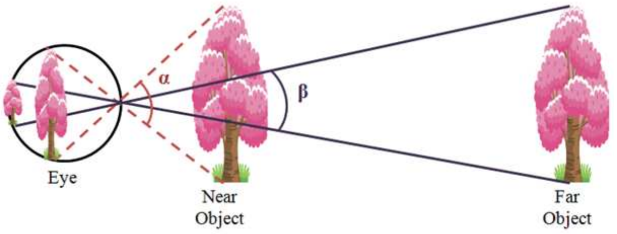


Fig.2: *Principal of Linear Perspective*

When watching the near object, we need the wider viewing angle (α) than the viewing angle (β) for the far object. As a result, the size of the projected near object on the retina is larger than that of the projected far object.

Therefore, lines that are parallel in the 3-D world appear to get closer together as they recede in the distance. The fact can help us to figure out the distance between two objects. This phenomenon is related to the relative size cue and vanishing effects. Because those often work together, linear perspective provides the strong effect for depth perception.

Relative Size. We use prior knowledge of the size of an object and combine that with information about the viewing angle to determine how far away from us an object is located.

If two objects are known to be the same size but their absolute size is unknown, relative size cues can provide information about their relative depth. If one subtends a larger visual angle on the retina than the other, the object which subtends the larger visual angle appears closer. Since the visual angle of an object projected onto the retina decreases with distance, this information can be combined with previous knowledge of the objects size to determine the absolute depth of the object. For example, we are generally familiar with the size of an average automobile. This prior knowledge can be combined with information about the angle which subtends on the retina to determine the absolute depth of an automobile.

Aerial Perspective. Aerial perspective refers to the effect that the atmosphere has on the appearance of an object as it is viewed from a distance. It explains that as the distance between an object and a viewer increases, the contrast between the object and its background decreases, and the contrast of any markings or details within the object also decreases. The colors of the object also become less saturated and shift towards the background color, which is usually blue, but under some conditions may be some other color.

Shadow. Shadows of the object can be a good element to estimate its depth. There are two types: attached or cast shadows. Attached shadows lie directly on the objects by whose shape, spatial orientation, and distance from the light source. Cast shadows are thrown from one object onto another, or from one part onto another of the same object [6].

Attached shadows are frequently the source of 'three-dimensionality' of the objects. For example, in shape from shading, the brightness of the surface is used to compute the 3-D shape of the surface [7]. Furthermore, artists commonly use attached shadows in pencil and ink sketches to induce a sense of 3-D surface shape. Thus, by exaggerating the bright and shaded part of the object, we can emphasize the volume of the object.

Focus. When seeing objects, we tend to only focus the object in the region of interest. The lens of the eye can change its shape to bring objects at different distances into focus. Knowing at what distance the lens is focused when viewing an object means knowing the approximate distance to that object.

Occlusion. Occlusion of objects by others is also a clue which provides information about relative distance. When the boundary of an object is interrupted by the presence of another object, we use this pattern of blocking as a cue to determine the object as more distant from us. The near object is as interposed between the far object and us. Since occlusion artificially generated for depth enhancement can seriously distort original images, we do not deal with this cue in this paper.

Motion Parallax. Motion parallax is the depth cue that results from motion of an observer. As an observer move, objects that are closer to us move farther across our field of view than do objects that are in the distance. If information about the direction and velocity of movement is known, motion parallax can provide absolute depth information.

Figure 3 shows the examples of monocular depth cues: focus effect, linear perspective, aerial perspective, and shadow effect.

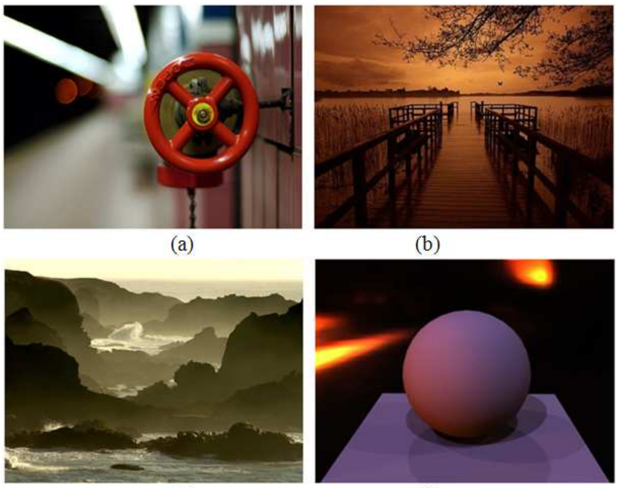


Fig.3: *Examples of Monocular Depth Cues: (a) Focus Effect, (b) Linear Perspective, (c) Aerial Perspective, and (d) Shadow Effect*

3. MONOCULAR DEPTH CUE ENHANCEMENT

In this paper, we propose the methods which enhance various monocular depth cues in a single-view image. It can provide improved depth perception on conventional 2-D display devices, especially in portable devices. Among the monocular depth cues introduced in the previous section, four main monocular cues are selected and enhanced. Those are listed in Fig. 4. Color transformation is the comprehensive meaning of aerial perspective and advancing and receding color theory. In our algorithm, two input images, color and its corresponding depth images, are used to produce the output image with enhanced depth cues.

On the Basis of the objects attracting interest, region of interest (ROI) is estimated. By considering ROI, we determine the kinds of depth cues to be enhanced and its degree. After color transformation and enhancement of shadow and linear perspective, we blur a background region so as to emphasize ROI. The process order is designed to minimize distortions of the output image.

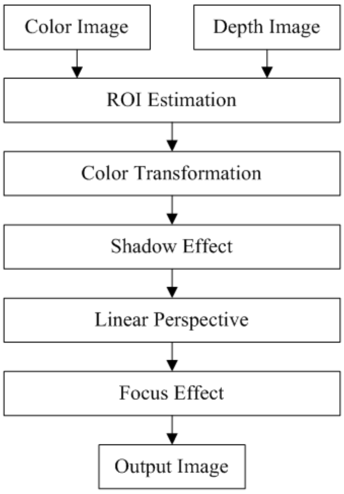


Fig.4: *Flowchart of Proposed Algorithm*

3.1 Aerial Perspective

Far objects will contrast less against each other, in terms of how light or dark they are, and will overall look less vibrantly-colored. The reason for this is scattering phenomenon due to small water and dust particles in the air between the eye and the object. The colors of objects at a great distance away also appear bluer.

We change the color of a background based on HSV domain using the concept of aerial perspective and advancing and receding color theory. On the basis of the scattering phenomenon, the intensity component of the background is modified. The perceived intensity consists of I_{att} and I_{air}. I_{att} is a light attenuation term expressed using Bouguer's law [8]. I_{air} is an air light caused by ambient light scattered towards the point of observation along the observation ray. (1) is the expression of the light scattering phenomenon caused by electromagnetic radiation.

$$I = I_{att} + I_{air} = I_0 e^{-\beta(\lambda)z} + I_\infty (1 - e^{-\beta(\lambda)z}) \quad (1)$$

where I_0 is the intensity that would be perceived in the absence of attenuation, $\beta(\lambda)$ is the extinction coefficient for a given wave-length λ, z is the distance between the camera and the scene area being imaged on the particular element and I_∞ is the intensity at the horizon. In this paper, we assume that $\beta(\lambda)$ is a constant value in the visible spectrum.

In the cases of hue and saturation, we change the values of each component of a background by using advancing and receding color theory. Advancing color is the set of colors that seem to come forward, and receding color is the set of colors that seem to move back slightly. Advancing colors have long wave-length and are highly saturated, while receding colors have short wave-length and are rarely saturated. In our approach, saturation and hue values are changed to have receding colors as the depth value decreases, and to have advancing colors as the depth value increases. Thus, we change each component of HSV values of a background as follows.

$$\begin{aligned} H' &= H + H_{max\_step}(1 - e^{-\beta z}) \\ S' &= S(1 - e^{-\beta z}) \\ V' &= V e^{\beta z} + V_{horizon}(1 - e^{-\beta z} \end{aligned} \quad (2)$$

where H', S', and V' are changed hue, saturation, and intensity values after color transformation, and H, S, and V are original hue, saturation, and intensity values. These equations are set based on Bouguer's law and advancing and receding theory.

Figure 5 shows the results of color transformation. The left column images are original images, and the right column images are result images. The background of the right column images looks receding compared with that of the left column images. On the other hand, the background of the left column images looks advancing compared with that of the right column images.

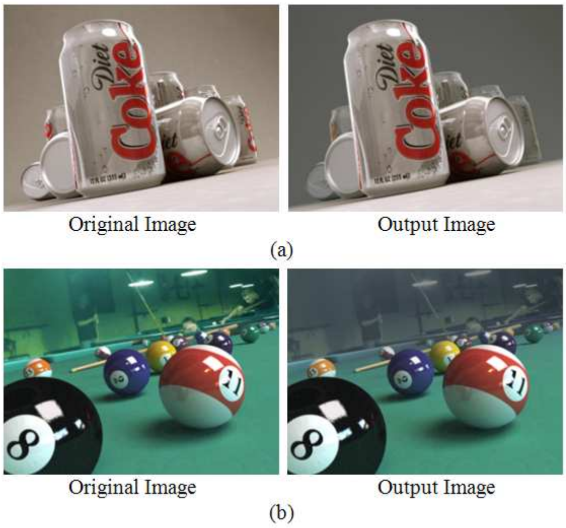


Fig.5: *Result of Color Transformation*

3.2 Shadow Effect

Shadow information of an object reflects the shape or volume of the object. The volume of the object can be emphasized by exploiting this shadow information. Thus, we propose the approach to produce ROI of single-view images with emphasized shadow.

The color space of ROI is transformed from RGB to HSV. The luminance component of ROI is modified by using the representative formula which is often used for gamma correction. Gamma correction is nonlinear operation, and defined by the following power-law expression, (3).

$$V'(i,j) = c\left(\frac{V(i,j)}{255}\right)^{\frac{1}{\gamma}} \times 255 \quad (3)$$

where c is the scaling factor, and r is the reciprocal number of gamma. The gamma correction nonlinearly changes the input values according to the gamma value. When the gamma value is smaller than one, dark regions become darker. In general, the gamma correction provides the similar effect with shadow emphasis, because shadow regions have low luminance values. This process is only applied to ROI.

Figure 6 shows the result of enhancing shadow effect. The left column images are the original images, and the right column images are the results after enhancing the shadow of ROI. The volume of ROI is emphasized, so it seems to be popped out.

3.3 Linear Perspective

In order to enhance linear perspective, we make the size of near objects larger and the size of far objects smaller, and move the objects to proper positions. The degrees of resizing and moving should be determined under physical calculation considering distances and viewing angles. In this paper, we employ 3-D warping technique with adaptive intrinsic camera parameters. A directional dilation and the in-paint methods are used to fill holes induced by the 3-D warping.

A. 3-D warping with Adaptive Intrinsic Camera Parameters

The 3-D warping [9] is a very popular technique to synthesize a virtual-viewpoint image with a current viewpoint image and its corresponding depth image.

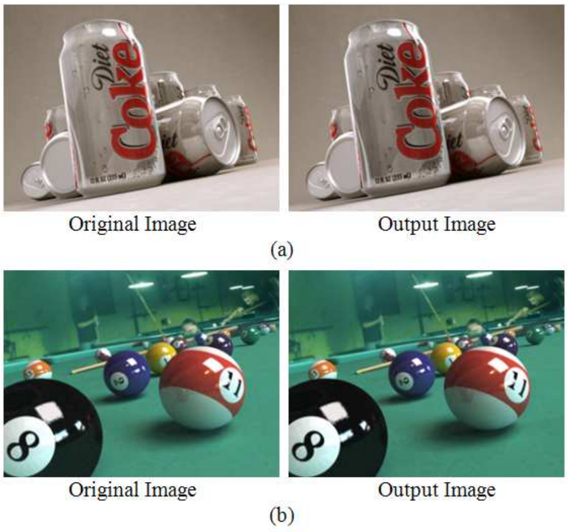


Fig.6: *Result of Shadow Effect*

In this case, extrinsic camera parameters are modified according to the position of the virtual-viewpoint. We apply the 3-D warping to get the enhanced linear perspective in this paper. Figure 7 shows the main algorithm of texture rearrangement using the 3-D warping. We virtually capture near and far objects with wider and narrower viewing angles, respectively.

The 3-D warping converts the coordinates of 2-D images to the 3-D world coordinates by (4),

$$(x, y, z)^T = \overline{RA}^{-1}(u, v, 1)^T d + \bar{t} \tag{4}$$

where d is distance between the camera and the pixel,$\bar{A}$ is the intrinsic camera parameter, and $\bar{R}$ and $\bar{t}$ are the extrinsic camera parameters. We can get the value of d from the corresponding depth image. The extrinsic camera parameters describe the orientation and position of the camera and explain the relationship between the world and camera coordinate systems.

The 3-D warping re-projects the objects in the world coordinates into the 2-D image coordinates by (5) which is the reversed form of (4).

$$\bar{A}\bar{R}^{-1}\{(x, y, z)^T - \bar{t}\} = (u, v, 1)^T \tag{5}$$

In both formulas, (4) and (5), the intrinsic parameter relates the camera's coordinate system to the idealized coordinate system and includes physical parameters such as the focal length of the lens and the size of the pixels [10]. It is possible to associate with a camera and a normalized image plane parallel to its physical retina but located at a unit distance from the pinhole. We attach to this plane its own coordinate system with an origin where the optical axis pierces it.

In general, the origin of the camera coordinate system is at a corner of the retina and not at its center, and the center of the CCD matrix usually does not coincide with the principal point. This adds two parameters that define the position of the principal point in the retinal coordinate system. The camera coordinate system may also be skewed due to some manufacturing error, so the angle between the two image axes is not equal to 90°.

The detail elements of the intrinsic parameter are shown in (6), and f, contained in a and b, relates to the focal length.

$$\bar{A} = \begin{bmatrix} a & -a\cot\theta & u_0 \\ 1 & b/\sin\theta & v_0 \\ 0 & 0 & 1 \end{bmatrix} \tag{6}$$
$$where, a = kf, b = lf$$

In (6), u_0 and v_0 define the position of the principal point in the retinal coordinate system, k and l represent pixel dimensions, and θ is the rotated angle.

$$p = \bar{A}\hat{p}, where\ p = \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} \tag{7}$$

Putting it all together, we obtain

$$p = \frac{1}{z}\bar{M}P, where\ \bar{M} = (\bar{A} \ \ 0) \tag{8}$$

and $P = (x, y, z, 1)^T$ denotes the homogeneous coordinate vector of P in the camera coordinate system. In other words, homogeneous coordinates can be used to represent the perspective projection.

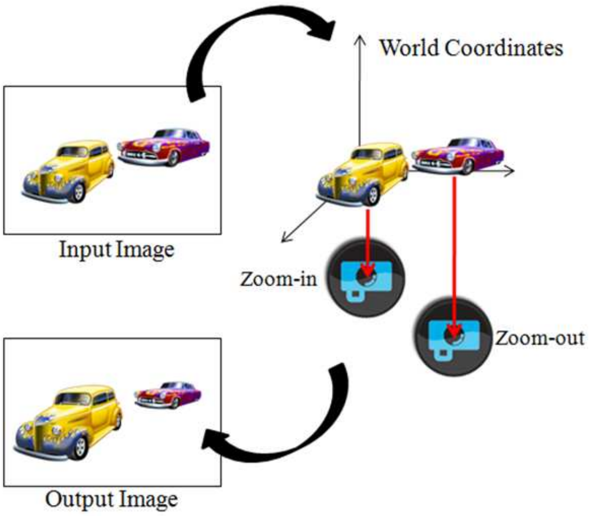


Fig.7: *3-D Warping with Adaptive Camera Parameters*

For our approach, we fix other parameters excepting the focal length. The large and small focal length induces zoom-in and zoom-out effects of camera. Hence we adjust the focal length parameter according to the depth value of objects and then project the objects to the 2-D image coordinates. The new focal length is calculated by (9).

$$f' = ((Depth_{ROI} - Depth)/Depth_{ROI})^{C_1}/C_2 + (1 - C_2^{-1}) \tag{9}$$

where DepthROI stands for the depth value of ROI, and Depth is current pixel's depth value. Two coefficients, C_1 and C_2, control the degree of rearrangement: C_1 is for non-linearity, and C_2 controls the range of rearrangement. A small C_2 coefficient induces large enhancing effect, but it distorts the original image. It is a trade-off between the degree of effects and distortions. Users can control two coefficients according to its applications and input images.

After this process, we can obtain the output image having enhanced linear perspective, but it contains holes induced by the 3-D warping. The hole is a newly exposed region due to the texture rearrangement, and the input image does not have any information for this region. Because the holes act like distortions, they should be filled with appropriate values.

B. Hole Filling using Directional Dilation

During the 3-D warping, the holes often are filled with zero values, thus it looks like huge shadow as shown in Fig. 8. Since these regions degrade the image quality, we should fill them with proper values by considering adjacent regions. In this paper, we use the in-paint technique to fill holes, which reconstructs small damaged region of an image. It refers to texture information of neighbor regions and restores the values in the holes. The in-paint method selects the region to be filled, and color information is propagated inward from the region boundaries. In order to produce a perceptually plausible reconstruction, it should attempt to continue strong features as smoothly as possible inside the reconstructed region.

Fig.8: *Holes Caused by 3D Warping*

The 3-D warping generates holes in the background region due to its smaller depth value, and both foreground and background often surround the hole region. Although the value for the holes should be estimated on the basis of only background's information, the in-paint algorithm refers to both foreground and background information, and recovers values with them. It induces serious distortions in the restored image as shown in Fig. 9(b). The dilated depth image can solve this distortion, since it makes the holes be surrounded with only background region. The dilation process for a depth image enlarges the size of foreground regions.

Although this approach can prevent above problem, it causes another problem that unnecessary background regions are attached to foreground region during 3-D warping. We propose therefore the directional dilation method for this problem. As a result of the 3-D warping with adaptive intrinsic camera parameter, the holes are located near to the foreground objects, especially to the direction toward the image center. Based on this property, we define 3×3 mask and the only apply dilation process to the neighbor pixels when the following conditions are satisfied.

1. *The depth value of the neighboring pixel is smaller than that of the current pixel.*
2. *The neighboring pixel lies on the extended line connecting the current pixel and the image center.*
3. *The distance between the neighboring pixel and the image center is shorter than that between the current pixel and the image center.*

With directional dilation considering the conditions, the regions located at the boundaries toward the image center are dilated. It makes the 3-D warping avoid unwanted attachments of background regions.

By applying the in-paint algorithm with both color and directionally dilated depth images, the clear output image can be obtained. The example is shown in Fig. 9(c). While the in-paint without the directional dilation induces the blended region of foreground and background as shown in Fig. 9(b), the directional dilation method selectively expands the foreground so as to minimize visual distortions.

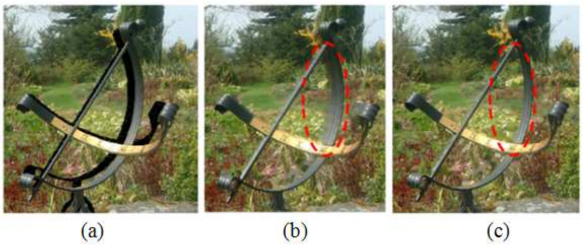
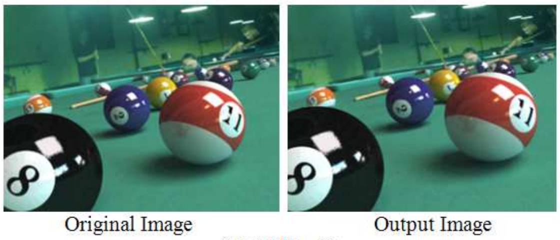


Fig.9: *Hole Filling: (a) Holes, and Hole Filling (b) without and (c) with Directional Dilation*

Figure 10 represents the results of enhancing linear perspective. The foreground sizes become larger than original ones, and there is no considerable distortions caused by holes.

(a) "Diet Coke"



(b) "Billiard"

Fig.10: *Result of Enhancing Shadow Effect*

3.4 Focus Effect

Generally, when a picture is taken, a foreground is focused and a background is blurred according to the distance from ROI. Thus, the blur amount can be a good measure for the depth [11]. We use this fact to enhance depth perception of the image. We apply the Gaussian filter to blur the background of the image. The amount of blur is calculated according the depth values of the background. The Gaussian filter is defined as follows.

$$G(x,y) = \frac{1}{2\pi\sigma^2} e \frac{-(x^2+y^2)}{2\sigma^2} \tag{10}$$

where x and y are pixel positions, and σ is the variance of the Gaussian function. We assume that ROI is found based on the depth value. We change the variance of Gaussian filter, σ , according to the depth values. For the farther background from ROI, the larger σ is defined. That is, if the region is farther away from ROI, then the region is blurred more. To avoid blurring on the boundary of ROI, we assign zeros to the region corresponding to ROI in Gaussian mask if the mask and ROI are overlapped as shown in Fig. 11.

Fig.11: *Gaussian Blurring for Neighbor Region of ROI*

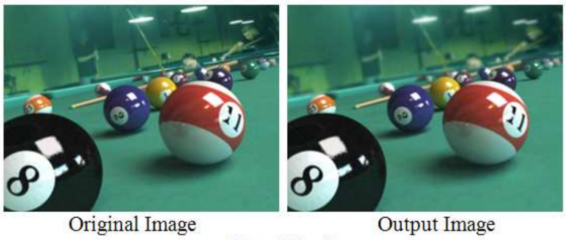
After blurring, the sharpening effect is applied to ROI to emphasize it. For sharpening effect of ROI, we use the following equation.

$$g(x,y) = f(x,y) - c\nabla^2 f(x,y) \tag{11}$$

where $g(x,y)$ is the sharpened result, $f(x,y)$ is the input image, and $\nabla^2$is a Laplacian operator. c can have different value according to the center of mask. When the center is positive or negative, it is set with -1 or 1, respectively. We use the following Laplacian operator.

$$\begin{aligned}\nabla^2 f(x,y) &= f(x+1,y) + f(x-1,y) + f(x,y+1) \\ &\quad + f(x,y-1) - 4f(x,y) \qquad (12)\end{aligned}$$

(a) "Diet Coke"



(b) "Billiard"

Fig.12: *Results of Enhancing Focus Effect*

Figure 12 shows the comparison of original image with the result of the blurring background and sharpening foreground according to the depth values. As can be seen in the figure, the image has the effect that the foreground corresponding to ROI is focused and the background is defocused. When comparing the output images with the original images, we can feel that the background in the right image seems to be further away than the original image, and the foreground seems to be closer than the original image.

4. EXPERIMENTAL RESULTS

All the enhancing techniques described above, linear perspective, color transformation, focus, and shadow effects, are integrated to generate the result image whose depth perception is enhanced. The test images are downloaded from the website [12]. Figure 13(a) and (b) show two input images, a color image and its corresponding depth image. The color image gives information on the volume of ROI conveyed with the shadow, and the depth image gives information on the location of ROI and depth values.

Figure 13(c) demonstrates the corresponding result images after whole procedures. The two balls in the image are emphasized, and the foreground region looks vivid and immersive.

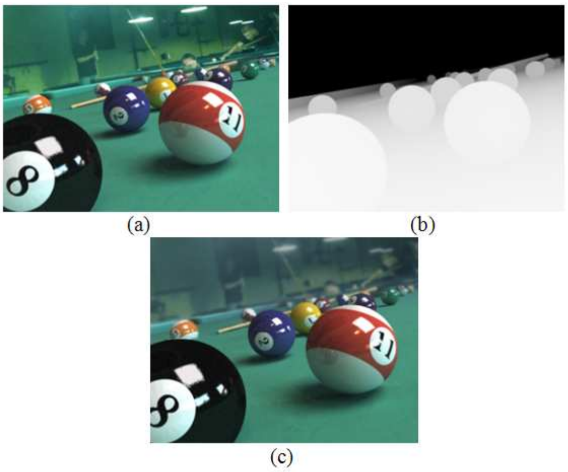

(a) (b) (c)

Fig.13: *Results for "Billiard" Image: (a) Color Image, (b) Depth Image, and (C) Output Image*

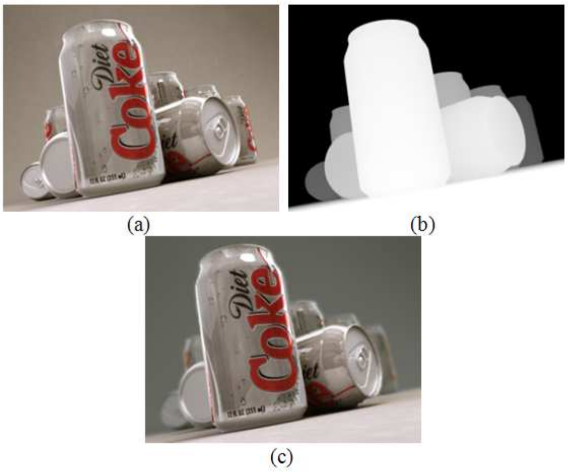

(a) (b) (c)

Fig.14: *Results for "Diet Coke" Image: (a) Color Image, (b) Depth Image, and (C) Output Image*

Figure 14, Fig. 15, and Fig. 16 show the results of other images. The depth perception of the image is improved, and the volume of ROI is emphasized when the result image is compared with the original image. The holes generated by enhancing linear perspective are successively filled with the proposed method, thus we cannot observe the serious distortions in the output images.

5. CONCLUSIONS

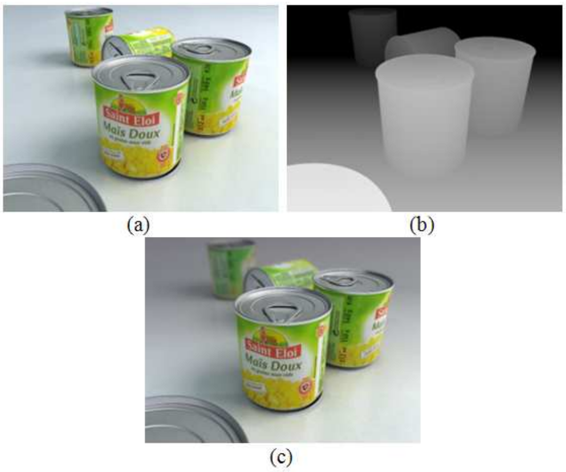

(a) (b) (c)

Fig.15: *Results for "Corn Box" Image: (a) Color Image, (b) Depth Image, and (C) Output Image*

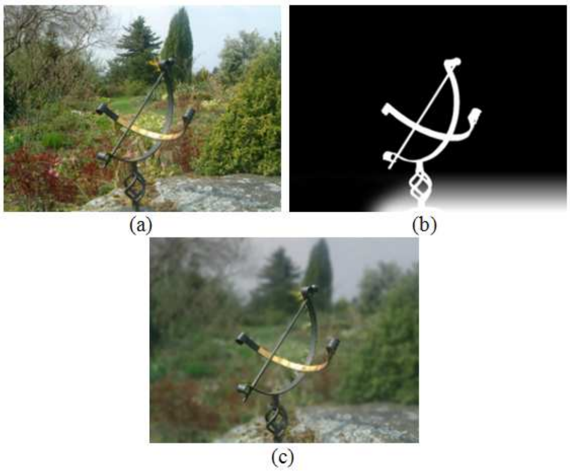

(a) (b) (c)

Fig.16: *Results for "Sundial" Image: (a) Color Image, (b) Depth Image, and (C) Output Image*

In this work, the methods to enhance depth perception of single-view images using a color image and its corresponding depth image are proposed. We adopt four enhancing techniques for aerial perspective, shadow effect, linear perspective, and focus effect to the original image to improve depth perception. We demonstrated that our algorithm gives good results that the volume of foreground is emphasized. The color of background is changed and blurred such that foreground looks conspicuous. The result image provides improved depth perception without considerable visual distortions. Our approach is very useful especially for the mobile display which has limitations for showing 3D images or videos.

6. ACKNOWLEDGEMENT

This research was supported by the MKE(The Ministry of Knowledge Economy), Korea, under the ITRC(Information Technology Research Center) support program supervised by the NIPA(National IT Industry Promotion Agency). (NIPA-2010-(C1090-

1011-0003))

References

[1] B. R. Cavoto and R. G. Cook, "The Contribution of Monocular Depth Cues to Scene Perception by Pigeons," *Psychological science: Journal of the American Psychological*, Vol. 17 No.7, pp.628-634, 2006.

[2] P. Kauff, N. Atzpadin, C. Fehn, M. Moller, O. Schreer, A. Smolic, and R. Tanger, "Depth map creation and image-based rendering for advanced 3DTV services providing interoperability and scalability," *Signal Processing Image Communication*, Vol. 22 No.2, pp.217-234, 2007.

[3] C. Fehn, R. Barre, and S. Pastoor, "Interactive 3-DTV-concepts and key technologies," *Proceedings of the IEEE, Special Issue on 3-D Technologies for Imaging & Display*, Vol. 94 No.3, pp.524-538, 2006.

[4] E. B. Goldstein, Sensation and perception, Pacific Grove CA, Wadsworth, 2002.

[5] E. Emmert, "Grossen verhalnisse der Nachbidder," *Klinische Monatsblätter für Augenheilkunde*, Vol. 19, 1881.

[6] P.S. Tsai and M. Shah, "Shape from shading using linear approximation," *Image and Vision computing*, Vol. 12, pp.487-498, 1994.

[7] E. Salvador and A. Cavallo, "Shadow identification and classification using invariant color model," *Proceedings of International Conference on Acoustics, Speech, and Signal Processing*, Vol. 3, pp.1545-1548, 2001.

[8] W. E. Knowles Middleton, Vision through the atmosphere, University of Toronto Press, 1958.

[9] A. Smolic and D. McCutchen, "3DAV exploration of video-based rendering technology in MPEG," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 14 No.3, pp.348-356, 2004.

[10] D. A. Forsyth and J. Ponce, Computer Vision, a modern approach, Prentice Hall, 2003.

[11] G. Surya and M. Subbarao, "Depth from defocus by changing camera aperture: a spatial domain approach." *Proceedings of IEEE International Conference of computer Vision and Pattern Recognition*, pp.61-67, 1993.

[12] DOF PRO, http://dofpro.com/.

[13] J. Jung, S. Seo, S. Lee, and Y. Ho, "Enhanced Linear Perspective using Adaptive Intrinsic Camera Parameters," *Proceedings of International Conference on Embedded Systems and Intelligent Technology*, pp.67(1-6), 2010.

[14] J. Lee, I. Shin, J. Moon, K. Kim, S. Lee, and Y. Ho, "Enhancement of Depth Perception for 2D Images," *Proceedings of International Conference on Embedded Systems and Intelligent Technology*, pp.68(1-6), 2010.

Jae-Il Jung received his B.S. degree in electronics and communications engineering from Hanyang University, Korea, in 2005, and M.S degree in information display engineering from Hanyang University, Korea, in 2007. He is pursuing a Ph.D. degree in the Department of Information and Communication at the Gwangju Institute of Science and Technology (GIST). His research interests include digital signal processing, video coding, 3D television, realistic broadcasting and Markov random field modeling.

Jong-Ho Lee

In-Yong Shin

Ji-Hee Moon

Yo-Sung Ho received both B.S. and M.S. degrees in electronic engineering from Seoul National University, Korea, in 1981 and 1983, respectively, and Ph.D. degree in Electrical and Computer Engineering from the University of California, Santa Barbara, in 1990. He joined the Electronics and Telecommunications Research Institute (ETRI), Korea, in 1983. From 1990 to 1993, he was with Philips Laboratories, Briarcliff Manor, New York, where he was involved in development of the advanced digital high-definition television (AD-HDTV) system. In 1993, he rejoined the technical staff of ETRI and was involved in development of the Korea direct broadcast satellite (DBS) digital television and high-definition television systems. Since 1995, he has been with the Gwangju Institute of Science and Technology (GIST), where he is currently a professor in the Information and Communications Department. His research interests include digital image and video coding, image analysis and image restoration, advanced coding techniques, digital video and audio broadcasting, 3D television, and realistic broadcasting.