54.2: Natural Stereo Depth Creation Methodology for a Real-time 2D-to-3D Image Conversion

T. Iinuma, H. Murata, S. Yamashita, K. Oyamada
Hypermedia Research Center, SANYO Electric Co., Ltd., Moriguchi, Osaka, Japan

Abstract

We propose a "Depth Creation by stereo Perception (DCP) " method that creates more natural and impressive 3D images than ever from input 2D images by their stereo perceptual information, which includes horizontal motion, color, contrast, sharpness, and composition. We also propose a "Linear Depth Control (LDC) " method that controls the binocular parallax of the created 3D images by their image depth position in order to improve the 3D effect of the converted images. And the real-time 2D-to-3D image conversion unit implemented these new methods realizes more natural 3D images corresponding to various 3D displays from 2D images of any types and scenes.

1. Introduction

Recently three-dimensional (3D) images have become more popular not only with the business use for amusement parks but also with the commercial use by CATV and Satellite TV. But 3D contents have not enough been provided to fill the demand for the consumer's market. Our estimation is that 3D business field would expand if 2D images could be converted into 3D images easily.

The 2D-to-3D image conversion technology using the "Modified Time Difference (MTD) " method had been developed in 1994. The MTD method allows to convert 2D images into 3D with binocular parallax by selecting those images which would be a stereo-pair according to the detected horizontal motion in the sequential input images. But there are certain images like still pictures that can't be converted well by this method.

Then the "Computed Image Depth (CID) " method which can convert any types of 2D images into 3D had been developed in 1997. The 3D image is generated with the CID method by computing the depth of each part of the input 2D image with its contrast, sharpness, and composition. Furthermore the adaptive control method between the MTD and the CID had been developed in order to improve the 3D effect of the converted images. But this method can't have more effect than either only the MTD or only the CID.

The "Depth Creation by stereo Perception (DCP) " method has been developed to solve this subject. The DCP can create more natural and impressive 3D images from input 2D images than the adaptive control method between the MTD and the CID. Furthermore, we propose a "Linear Depth Control (LDC) " method that controls the binocular parallax of the created 3D images by their image depth position in order to improve the 3D effect of the converted images. These techniques have been implemented into the real-time 2D-to-3D image conversion unit.

2. Depth Creation by stereo Perception

The DCP method creates 3D images with binocular parallax at any part of the images by the integrated method of the MTD and the CID. The created 3D images have the definite binocular parallax at any time and scene by the integrated stereo perceptual information.

When the input 2D images have the object with the complicated motions or no motions, the DCP is the same behavior as the CID. At first, this method calculates the relative depth information by the perspective according to their color, contrast, sharpness, and composition of each part of the input 2D image. Secondly, this method generates the left and right eye images according to the depth information by shifting each area of the input image.

When the input images have simple horizontal motion, the DCP is the integrated behavior of the MTD and the CID. At first, this method calculates the integrated depth information that can be obtained from the horizontal motion and the perspective according to their color, contrast, sharpness, and composition. Secondly, this method generates the 3D images according to selecting the stereo-pair images from the input sequential images by the horizontal motion, and shifting the left and right images by the integrated depth information.

Figure 1 shows that each method, the MTD, the CID, and the DCP, creates the 3D images from the 2D images that the bus is running from the left to the right.

The MTD can not create binocular parallax at the background without motion though the relative binocular parallax between the bus as the moving area and the background as the still area is generated. But the bus as the moving area in the created 3D images has a stereo-occlusion. The CID can not create binocular parallax with stereo-occlusion though the binocular parallax is generated at any area of the images. The DCP can create binocular parallax at any area with stereo-occlusion at the bus as the moving area.

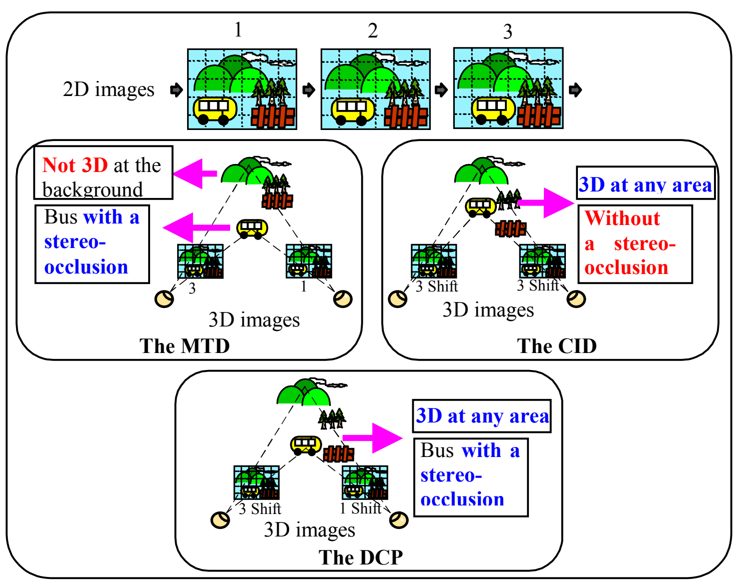

Figure 1 The effect of each method

ISSN0000-0966X/00/3101-1212-$1.00+.00

Figure 2 shows the basic principle of the DCP method. This method consists of the following four processes.

First, Detecting and Grouping process detects the stereo perceptual information of the input images and groups the areas according to the detecting information, which contains the horizontal motion, color, sharpness, and contrast value. For example, the adjacent areas that have close color and motion are grouped. The following processes, which are Occlusion and Perspective process and Depth Parameter process, work based on these grouped areas.

Second, Occlusion and Perspective process calculates the time difference between left and right images of Occlusion Images according to the horizontal motion, and calculates Occlusion Parameter and Perspective Parameter. At first, the time difference is calculated by the horizontal motion of each area according to the same behavior as the MTD. Next, Occlusion Parameter, that is the stereo perceptual information according to binocular parallax with stereo-occlusion by the calculated time difference and the average motion of each area, is calculated. Perspective Parameter, that is the relative depth information by the perspective of each area according to its sharpness, contrast, and image composition, is calculated. Generally in the photographs and the TV images, the near-positioned objects have higher sharpness and higher contrast than the far-positioned objects or the background images. And the image composition has the tendency that the center or the bottom side of images is nearer than the upper side. So, Perspective Parameter of each area is decided by the average of each area's sharpness and contrast value that is weighted by the image composition.

Third, Depth Parameter process calculates Depth Parameter from Occlusion Parameter and Perspective Parameter, and calculates the binocular parallax of each area to create the 3D images according to Depth Parameter. The each area's binocular parallax that calculated in this process decides the 3D effect of the converted images. The range of the binocular parallax between the nearest area and the farthest area in the converted 3D image can be changed by user's preference.

Fourth, Image Generation process selects Occlusion Images by the calculated time difference between left and right images of the second process, and shifts each area of the left and right images according to the calculated binocular parallax of each area by Depth Parameter of the third process. In the scene that the bus is running from the left to the right, the present image is given to the left eye and the delay image is given to the right eye. And, the bus can be seen as if it were popping out of the background. In the near area, the left eye's image is made by shifting the input images to the right, and the right eye's image is made by shifting to the left. In the far area, the both images are made by shifting to vice versa. Then, the 3D images are generated to have the binocular parallax calculated by Depth Parameter. Each shift value is adjusted to decrease the steep change of the binocular parallax between the adjacent areas.

As a result, the 3D images are created within the definite depth position at any part of the images, any time and scene, by generating binocular parallax from the perspective in the still area of the moving images, with stereo-occlusion by the horizontal motion, and from the integrated stereo perceptual information. So the natural and impressive 3D images are generated by the DCP.

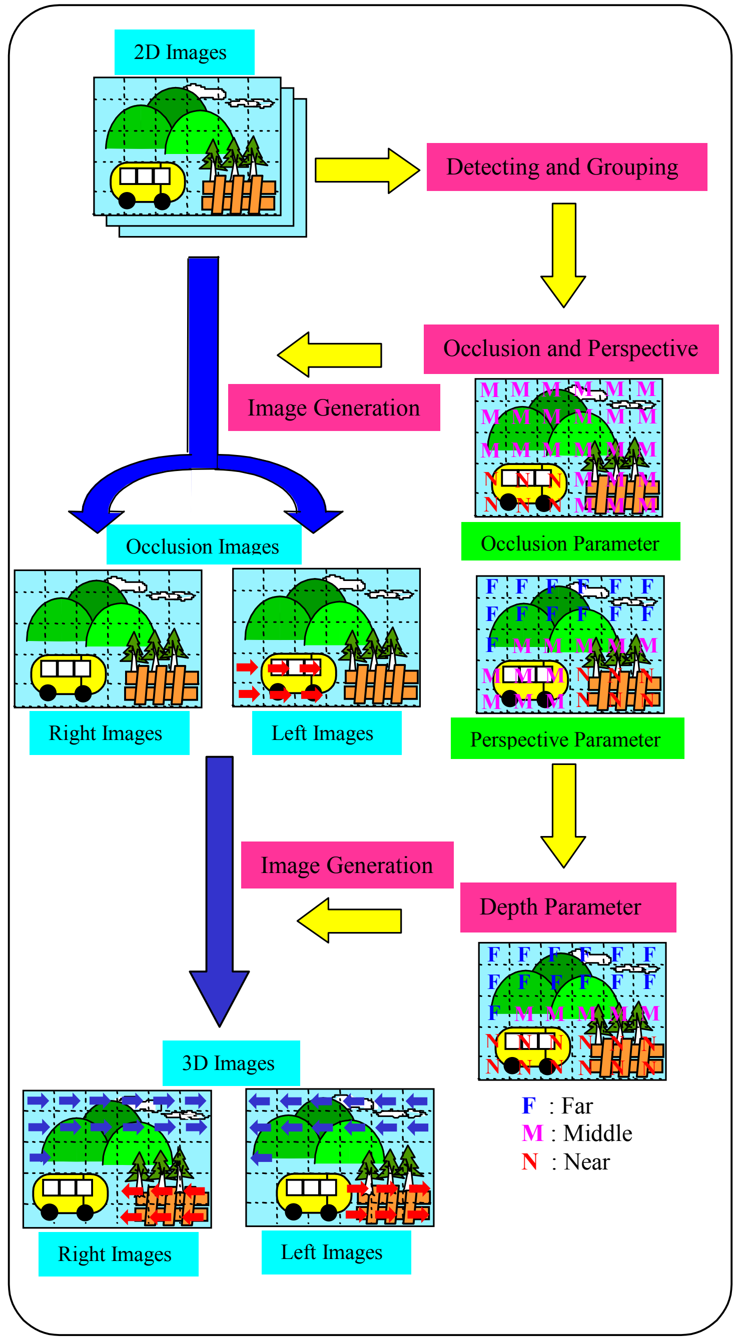

Figure 2 Principle of the DCP

3. Linear Depth Control

The "Linear Depth Control (LDC)" method controls the binocular parallax of the created 3D images by their image depth position in order to improve the 3D effect of the converted images.

And the LDC can also control the area ratio between the front side and the rear side of the created 3D images regardless of the range of the depth position between the nearest and the farthest that can be varied by user's preference in order to generate impressive 3D images.

Figure 3 shows the relation between the depth position and the binocular parallax of the 3D images. The relation between the depth position and the binocular parallax is given by

$$Y = XD / (E - X) \quad (1)$$

where Y is the depth position of the 3D images, X is the binocular parallax, E is the distance of between each eye, D is the distance of between the display and the observer.

In the conventional method, the binocular parallax is decided proportionally to the calculated Depth Parameter. The relation between the binocular parallax and Depth Parameter is given by

$$X = AP \quad (2)$$

where P is the Depth Parameter, A is a coefficient of proportion.

According to (1) and (2), the relation between the depth position and Depth Parameter is given by

$$Y = ADP / (E - AP) \quad (3)$$

Figure 4 shows the relation between the depth position and Depth Parameter by the equation (3). This method doesn't realize proportional depth position of the created 3D images, that is, the depth position changes non-linear as the increase of Depth Parameter. The perceptible 3D effect by the depth position becomes weak in the front side of the screen, and becomes emphasized in the rear side.

The LDC method controls the binocular parallax, as the depth position becomes proportional to Depth Parameter. The relation between the depth position and Depth Parameter of the LDC is given by

$$Y = AP \quad (4)$$

Figure 5 shows the relation between the depth position and Depth Parameter by the equation (4).

The perceptible 3D effect by the depth position becomes linear by the LDC. As a result, more natural 3D images without the distortion are created.

Figure 6(b) shows the variance of the rear depth position's range without the change of the area ratio by the LDC, figure 6(a) shows the variance of the rear depth position's range with the change of the area ratio. In the case (a), the front side area of the created 3D image decreases according to the increase of the rear depth position's range. In the case (b), the LDC can vary the range of the depth position without the change of the area ratio.

And the LCD method can vary the area ratio without the change of the depth position's range too.

As a result, the area ratio and the range of the depth position of the created 3D images can be change freely, and more impressive 3D images can be created.

4. 2D-to-3D Conversion Unit

We developed the automatic and real-time 2D-to-3D image conversion unit adapting the newly proposed methods. Figure 7 shows the appearance and its applications. Figure 7(a) shows the

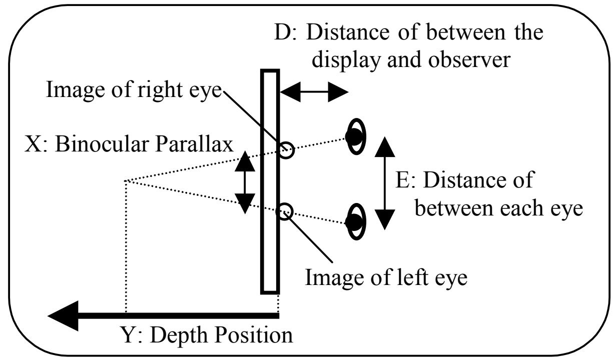

Figure 3 Relation between the binocular parallax and the depth position

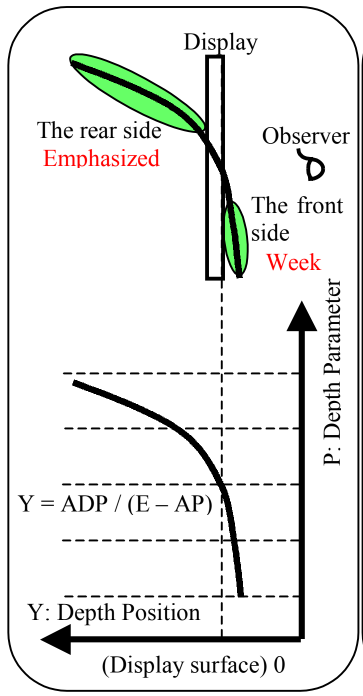

Figure 4 Relation between the depth position and the Depth Parameter

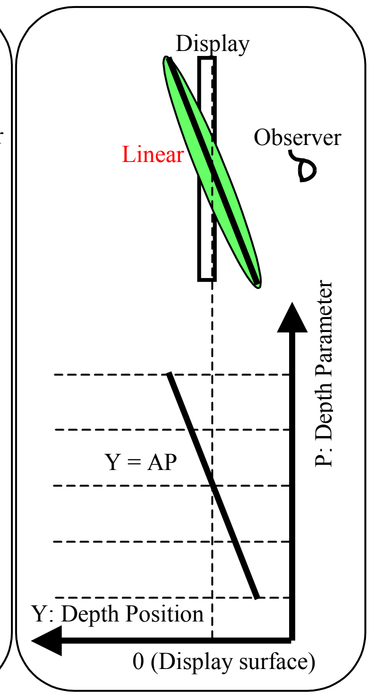

Figure 5 Relation between the depth position and the Depth Parameter (The LDC method)

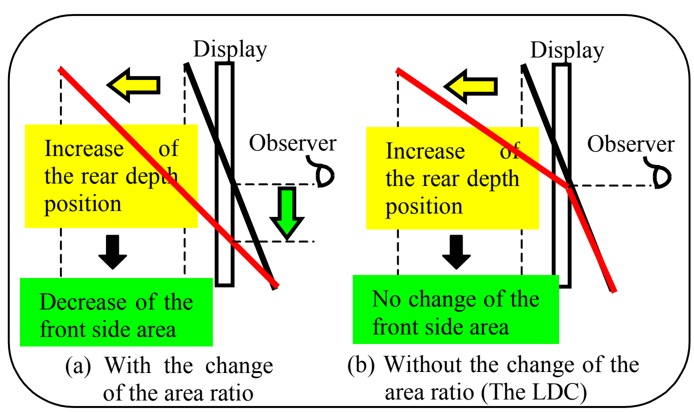

Figure 6 Relation between the depth position and the area ratio

appearance of the NTSC or PAL type, figure 7(b) shows the appearance of the HDTV or VGA type. This unit is able to convert any kind of 2D images into 3D images. Table 1 shows the specifications of the conversion unit of the NTSC type.

This unit has three output modes according to the employed 3D displays. The first is the field sequential output mode with normal scan rate (60Hz) applicable to the normal TV and some Head Mounted Displays. The second is the two channel of L and R image output mode applicable to the 3D theater for business and the glass-less 3D displays. The third is the RGB field sequential output mode with double scan rate (120Hz) applicable to the multi-scan PC displays.

5. Conclusion

These newly developed conversion methods realize to convert any types of visual sources into natural and realistic 3D images. Furthermore the 2D-to-3D conversion units implemented these methods can output 3D images according to various 3D displays from various input images, not only TV signals like NTSC, PAL, or HDTV standards but also PC graphics like VGA. These methods and conversion units would be providing more 3D programs, and establishing 3D business for consumer and entertainment fields.

Table 1 Specifications of the conversion unit

Input	Video input (composite) S-Video input (Y/C) Analog RGB (15.73kHz, 60Hz)
Output	Video 2ch output (composite) S-Video 2ch output (Y/C) Analog RBG (31.47kHz, 120Hz)
Power Source	100-120V AC, 50/60 Hz
Power Consumption	Approx. 10W
Weight	Approx. 5kg
Dimensions	430(W) x 44(H) x 370(D) mm

6. References

[1] H.Murata et al.: "Conversion of Two-Dimensional Images to Three Dimensions", SID Digest of Technical Papers, 39.4, pp859-862 (1995)

[2] T.Okino et al.: "New Television with 2D/3D image conversion technologies", SPIE Vol. 2653, pp96-103 (1996)

[3] H.Murata et al.: "A Real-Time 2-D to 3-D Image Conversion Technique Using Computed Image Depth", SID Digest of Technical Papers, 32.2, pp919-922 (1998)

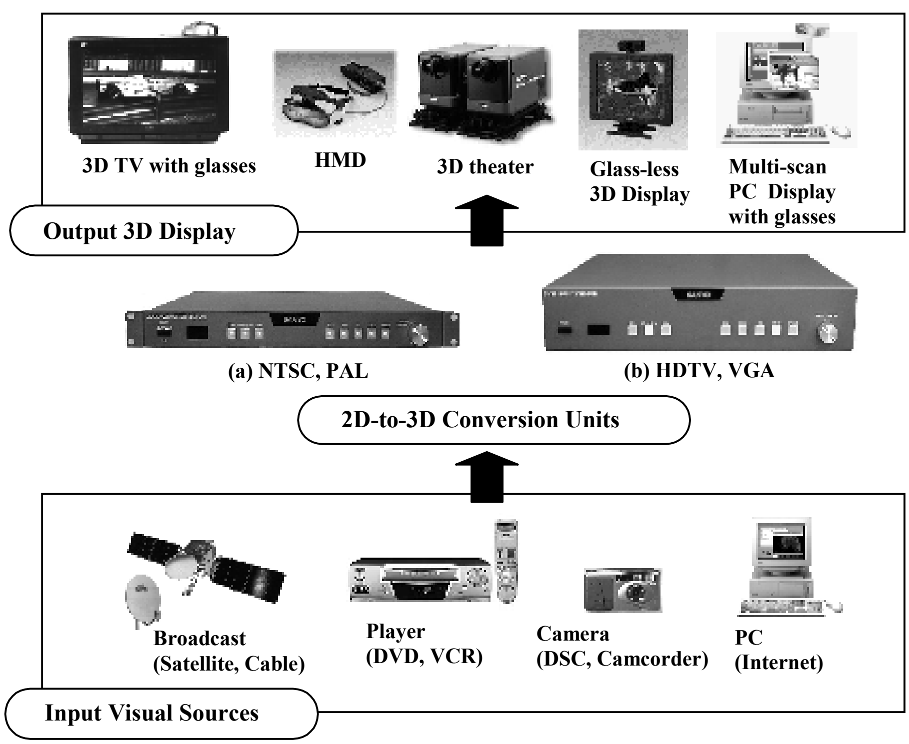

Figure 7 2D-to-3D Conversion Units and the applications