Stereoscopic Conversion of Monoscopic Video using Edge Direction Histogram

Jee-Hong Kim, Dong-Wook Kim**, Ji-Sang Yoo**

Dept. Electronic Engineering*, Dept. of Electronic Materials Engineering**.
Kwangwoon University
Seoul, Korea
E-mail: jjilong@kw.ac.kr, dwkim@kw.ac.kr, jsyoo@kw.ac.kr

ABSTRACT

This paper proposes an algorithm for creating stereoscopic video from a monoscopic video. A viewer uses depth perception clues called a vanishing point which is the farthest from a viewer's viewpoint in order to perceive depth information from objects and surroundings thereof to the viewer. The viewer estimates the vanishing point with geometrical features in monoscopic images, and can perceive the depth information with the relationship between the position of the vanishing point and the viewer's viewpoint. In this paper, we propose a method to estimate a vanishing point with edge direction histogram in a general monoscopic image and to create a depth map depending on the position of the vanishing point. With the conversion method proposed through the experiment results, it is seen that stable stereoscopic conversion of a given monoscopic video is achieved..

Keywords: stereoscopic image conversion, depth information

1. Introduction

Since our two eyes are apart by a given distance when we see an object, each slightly different image is captured in the left and the right eyes. This phenomenon is called binocular disparity. Our brain generalizes the two images of binocular disparity to recognize them as a synthesized image and we can have depth feeling. Conventional stereoscopic images are obtained with stereo cameras or image editing tools. However, generating such conventional stereoscopic images with a stereoscopic image input device is not cost-effective. Editing by means of image editing tools is time-consuming work, so that it is impossible to convert huge amount of images produced originally in monoscopic format into stereoscopic images. Therefore, if an efficient method can be found for stereoscopic conversion of monoscopic images, it will be very helpful to produce even more various contents. Since 1990, many researches have thus been done for finding a good method for converting such monoscopic images to stereoscopic images.

However, the conventional methods exhibit lowered quality in images by image variations. Also, the speed and directions of moving objects should be determined and we should also find out whether the movement is vertical or horizontal.

Images comprising all kinds of objects present in a room or outside consist of geometrical elements including straight lines and many parallel or orthogonal straight lines. Parallel straight lines in a 3D space get narrower as they are farther from the perspective images on a 2D plane and finally meet at one point that is called a vanishing point. The straight lines meeting at the vanishing point are also called vanishing lines. The vanishing point is one of depth clues used for recovering 3D spatial structure from 2D images. In particular, the vanishing point is a very important depth clue in outdoor taken images having distinctive lines in their background or in the images for corridors or indoor structures. Therefore, it is possible to estimate relative depth with the vanishing points.

In this paper, Chapter 2 describes the overall system of the proposed algorithm and in Chapter 3, a default depth map depending on the positions of vanishing points in images is described. In Chapter 4, described is the process for selecting the default depth map with the edge direction histogram. In Chapter 5, described is the process for creating the final depth map. In Chapter 6, the experiment results are shown and finally In Chapter 7, conclusion and future direction of study is represented.

2. System structure of the propose algorithm

Fig.4 shows the system block diagram for creating a depth map from 2D images, using edge direction histogram. The system extracts edges from input images and thins the

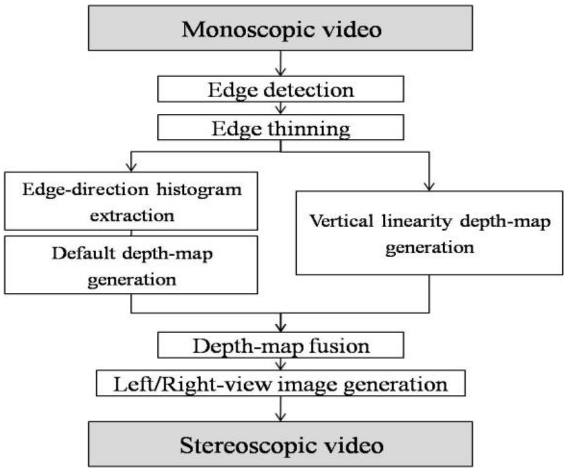


Figure 4. System block diagram of creating depth map with edge direction histogram

edges in order to obtain more exact edge direction histogram. Thereafter, the system obtains edge direction histogram using the thinned edges. From this histogram, it selects a default depth map, and synthesizes the depth map with the default depth map by using vertical boundary information between objects to create a final depth map. With the depth map obtained as described above, it produces right and left images, and synthesizes the right and left images to create 3D stereoscopic images..

3. Creating a default depth map

Since a vanishing point is a point where extended lines of straight lines meet, it represents the farthest point from a camera position by which the image concerned is captured. On the contrary, the point present at a position opposite to the vanishing point in an image indicates that it is located at the nearest distance. Therefore, it is possible to represent a relative depth in images with the vanishing points.
The vanishing point can be located at any place in an image depending on the geometrical features of the structure in the image. However, since the vanishing point in an image does not exist on the lower part of the image, 3 default depth maps are created, for the case the vanishing point is located on the left of the image, for the case the vanishing point is located on the upper part of the image, for the case the vanishing point is located on the right of the image. One more default depth map is created by assuming that objects are located on the lower part of the image. For the default depth maps created by right and left vanishing points, they should be upper left and upper right in an image in order to effectively keep the vertical depth feeling that is a basic assumption for creating a relative depth map to be described later in this paper. The case a vanishing point is located on the lower part of the image runs contrary to the vertical depth feeling described above and will thus not be further described in this paper.

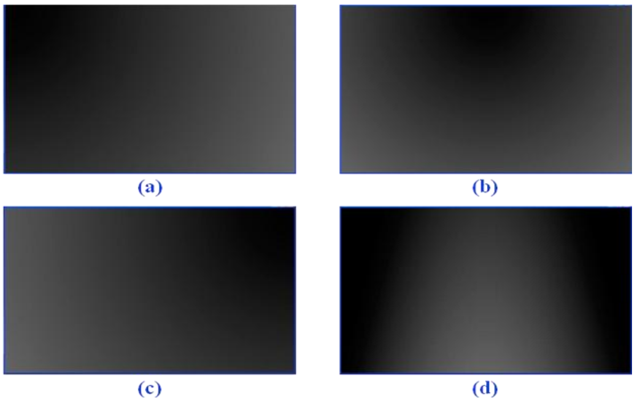


Figure 5. Default depth map images
(a)The vanishing point positioned on the left of the image.
(b)The vanishing point positioned on the upper part of the image.
(c) The vanishing point positioned on the right of the image.
(d) The object positioned on the lower center of the image.

4. Decision of vanishing points with the edge direction histogram

4.1 Edge thinning process

-1	0	1
-2	0	2
-1	0	1

1	2	1
0	0	0
-1	-2	-1

Figure 6. Sobel edge mask

Fig.6 shows Sobel mask. Sobel mask is used to detect edges from the input images. We then apply the thinning process to the extracted edges. Thinning process enhances accuracy in the next process for obtaining edge direction histogram and makes a vertical linearity depth map more reliable and robust to nois

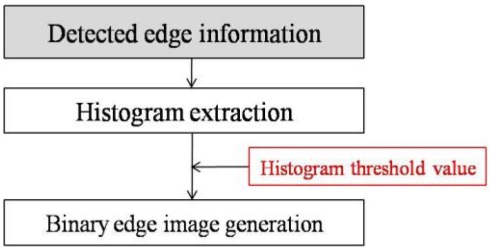
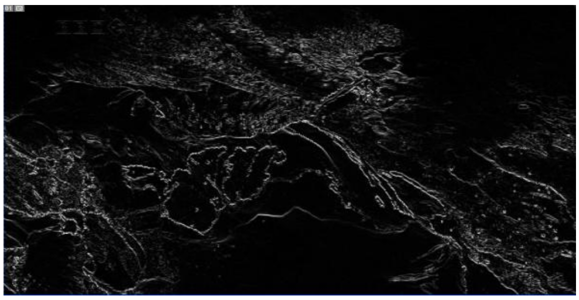


Figure 7. Edge thinning

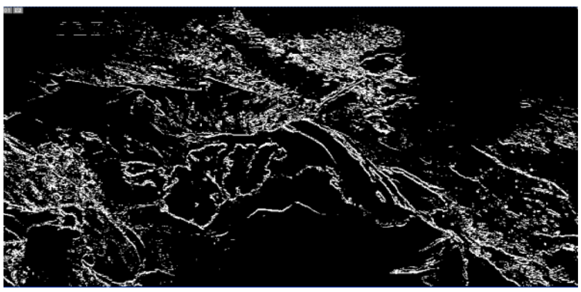
Fig.7 shows how edges are thinned. First we extract the histogram of detected edges by Sobel mask that has values from 0 to 255. After obtaining the histogram of detected edges, a binary edge map is obtained by using a threshold value in Eq. 1. In other words, the cumulative number of pixels up to some edge value starting from edge value of 0 is less than the threshold value, then we make all the pixels counted up to that edge value have value of 0 and the rest of pixels have 255 edge value.

$$threshold = num\_pixel \times rate \quad (1)$$

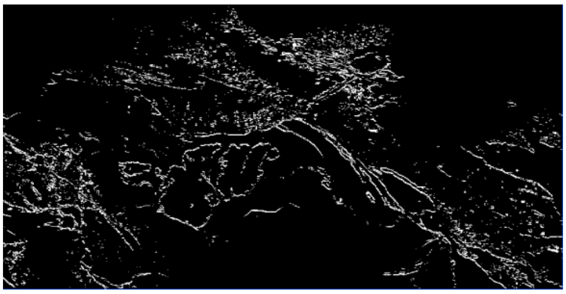
Where num_pixel represents the total number of pixels in an image and rate is determined experimentally



(a)



(b)



(c)

Figure 8. Edge thinning process
(a) Original edge map obtained by Sobel mask
(b) Binary edge map with rate=0.9 in Eq. (1),
(c) Binary edge map with rate=0.96 in Eq. (1).

Fig. 8(a), (b) and (c) show the original edge map and obtained binary edge maps with different values of rate through the thinning process, respectively. It is shown that Fig.8(c) has less edge noise and thinned edge components compared with Fig. 8(b) and can be more efficiently used in the next process

4.2 Extraction of edge direction histogram

From the obtained binary edge map in the previous subsection, now edge direction histogram can be obtained using Sobel edge mask again. Eq. (2) shows the x and y components of edge values when Sobel mask is applied to the binary edge map.

$$dx_{i,j} = p_{i-1,j+1} + 2\times p_{i,j+1} + p_{i+1,j+1} - p_{i-1,j-1} - 2\times p_{i,j-1} - p_{i+1,j-1}$$
$$dy_{i,j} = p_{i+1,j-1} + 2\times p_{i+1,j} + p_{i+1,j+1} - p_{i-1,j-1} - 2\times p_{i-1,j} - p_{i-1,j+1} \quad (2)$$

Where pi,j represents an edge value(0 or 255) of a pixel on the i-th column and j-th row and, dxi,j and dyi,j represents the x and y edge components of Sobel mask, respectively. Therefore, the direction of edge can be found as in Eq. (3).[5]

$$Ang(i,j) = \frac{180^\circ}{\pi} \times \arctan\left(\frac{dy_{i,j}}{dx_{i,j}}\right) \qquad |Ang(i,j)| < 90^\circ \quad (3)$$

Now each pixel is assigned an edge direction value that is in the range of from -90 ¡Æ to +90 ¡Æ and belongs to one of the 5 ranges as in Fig. 9 and it results in an edge direction histogram finally. The 5 ranges are specified by equally dividing them from -90 ¡Æ to +90 ¡Æ

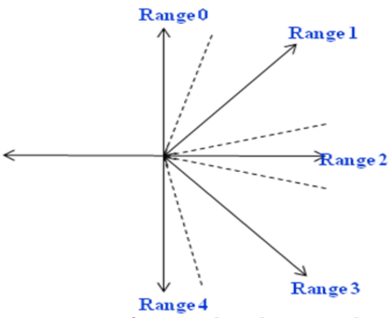


Figure 9. Ranges for an edge direction histogram

Now, one of the default depth map is decided by finding the position of vanishing point from the extracted edge direction histogram. First we estimate the number of pixels in each range and the range with the most number of pixels is selected and each selected range gives the corresponding default depth map as in table 1

Table 1. Decision of default depth map depending on ranges of edge direction histogram

Range	Default depth map
Range0	Depth map where the vanishing point is positioned on the upper side of an image
Range1	Depth map where the vanishing point is positioned on the right side of an image
Range2	Depth map where an object is positioned in the lower center of an image
Range3	Depth map where the vanishing point is positioned on the left side of an image
Range4	Depth map where an object is positioned in the lower center of the image

4.3 Generation of depth map depending on vertical position

Depth feeling depending on vertical positions in an image that is usually called a vertical linearity depth comes from the general fact that lower parts of an image is close from a viewer and the upper part thereof is relatively far in general. However, if a depth map is created only with a vertical linearity depth in an image, the created stereo images are inclined and cannot feel a difference of depth feeling between different objects. To compensate this problem, we use boundary information between objects.

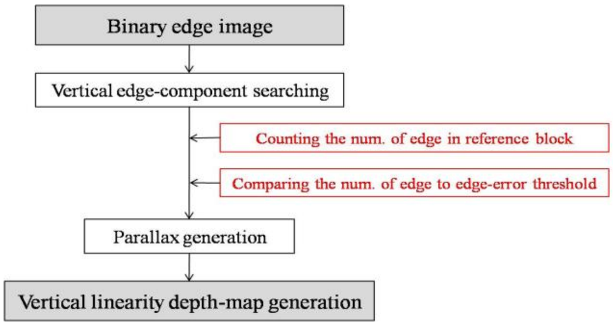


Figure 10. Creating a depth map depending on vertical position

Each object has its boundary. With a parallax difference in boundaries, we feel different depth feeling between different objects. Here we create a depth map depending on vertical boundaries robust to noise by using a binary edge map created through the thinning process explained in section 4.1. In this case, we have to consider pixels around a vertical boundary in order to reduce errors due to noise(isolated pixels) on the vertical boundary, that can cause totally different depth even though they have the same depth value.

5. Final depth map generation

The final depth map is created by using both the default depth map selected with an edge direction histogram and the vertical linearity depth map in the ratio of α to 1-α as in Eq. (4).

$$final\_depth(i,j) = \alpha \times vertical\_linearity(i,j) + (1-\alpha) \times default\_depth(i,j) \quad (4)$$

In Eq. (4), vertical_linearity(i,j) is a vertical linearity depth value at the position of (x, y), default_depth(i,j) is a determined default depth value at the same position, and flinal_depth(i,j) is the final depth value.

6. Experimental results

In this paper, we take 2600 frames of the documentary program “The World of Pure White Ice” by BBC with the size of 1280×720 to use them as a test video. The right and left images are stereoscopically synthesized in the interlaced manner by using the final depth map created with the proposed algorithm in this paper.

It turns out that among total 2600 frames, it was 9% for the case of using the default depth map in which the vanishing point was positioned on the left of the image, 4% for the case of using the default depth map in which the vanishing point was positioned on the right of the image, and 87% for the case of using the default depth map in which the object was positioned in the lower center of the image.

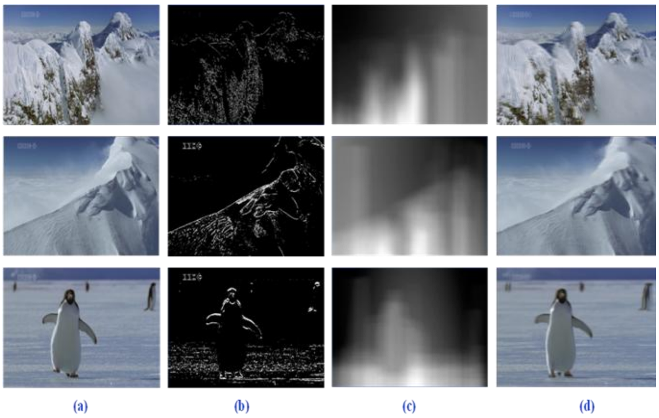


Figure 11. Result mages. From the top row to bottom row, with vanishing point on the left, right and with object on the lower center respectively.
(a)Original image (b) Edge map detected with Sobel mask (c) Final depth map (d) Synthesized stereoscopic image

Fig. 11 shows the result images. The images on the top row are for the case of using the default depth map in which the vanishing point was positioned on the left of the image. The images on the middle row are for the case of using the default depth map in which the vanishing point was positioned on the right of the image, and finally the images on the bottom row are for the case of using the default depth map in which the object was positioned in the lower center of the image.

In the proposed algorithm, we used the directional information of edges in an image and also vertical boundary information in an images, effective stereoscopic image conversion was able to be carried out even in images such as ones captured at a far distance, landscape or panoramic photographs. Experimental results show that the proposed algorithm has some advantages that conversion for images with high resolution can be done more rapidly comparing with conventional algorithms, and can be applied to all kinds of images. In the proposed algorithm, we could also solve the problem of significantly lowered accuracy when distinctive straight line components are not included in an image that is a week point in conventional algorithms[6],[7] using vanishing points for creating a depth map.

7. Conclusion

In this paper we proposed an algorithm for creating stereoscopic video from a monoscopic video.

Through experimental results, the proposed algorithm showed a very good performance with less computational complexity comparing with conventional algorithms for conversion. Also the proposed algorithm showed that it works well on any kind of images.

However, there is still difficulty in obtaining an edge direction histogram and creating a vertical linearity depth map for images of much noise or unclear images. Therefore, it is considered that further study is needed for diversifying the types of default depth maps, improving accuracy of deciding default depth maps.

Acknowledgement

This work was supported by grant No. (R01-2006-000-10199-0) from the Basic Research Program of the Korea Science & Engineering Foundation.

Reference

[1] T. Okino and et. al, "New television with 2D/3D image conversion techniques," SPIE, vol. 2653, Photonic West,1995

[2] B. J. Garcia. "Approaches to stereoscopic video based on spatial-temporal interpolation," SPIE, vol. 2635, Photonic West, 1990

[3] Y. Matsumoto, H. Terasaki, K Sugimoto and T. Arakawa, “Conversion system of monocular image sequence to stereo using motion parallax”, SPIE Vol. 3012, Photonic West, 1997.

[4] Ross, J., "Stereopsis by binocular delay", In nature, vol. 248, PP.354-364, 1974

[5] Feng PAN, Xiao LIN, Rahardja SUSANTO, Keng Pang LIM, Zheng Guo LI, Ge Nan FENG, Da Jun WU, and Si WU "Fast Mode Decision for Intra Prediction", JVT-G013, 7th meeting, Pattaya II, Thailand, 7-14 March, 2003

[6] S. Battiato, S. Curti, M. LaCascia, E. Scordato, M. Tortora, "Depth-Map Generation by Image Classification", In Proceedings of SPIE Electronic Imaging 2004, Three-Dimensional Image Capture and Applications VI, vol. 5302-13, San Jose, California, USA, January 2004

[7] Jongwon Han, Jinsu Jo, Yillbyoung Lee, “Relative depth-map generation from monocular image using a valid vanishing point”, The journal of Korea computer congress, vol.33, No.2(B), 2006