3D Perception Adjustment of Stereoscopic Images Based upon Depth Map

Jong In Gil, Seung Eun Jang, and Manbae Kim

Department of Computer and Communications Engineering,
Kangwon National University,
Chunchon, 200-701, Repubic of Korea
{jigil,jse4485,manbae}@kangwon.ac.kr

Abstract. Recently, a variety of stereoscopic contents have been provided to academic and industrial fields for broadcasting, movies and mobile materials. However, few works have been interested in the adjustment of 3D contents for diverse displays. For instance, movie contents suited to large screen frequently do not deliver the same 3D perception to small-size screen such as mobile phone, tabular PCs, etc. For this, this paper presents an adjustment method of stereoscopic contents. 2D+Depth is one of popular methods with which stereoscopic images are generated. For this, depth planes are derived based on a depth histogram. By adjusting depth planes, a new depth map is made. Then 2D+Depth produces a stereoscopic image. Experiments performed on various 2D+Depth images validate that the proposed methods deliver more enhanced 3D depth based on subjective evaluation experiments.

Keywords: stereoscopic perception, depth map adjustment, subjective test.

1 Introduction

The advances in stereoscopic video technologies have led to an increasing interest in various 3D applications [1, 2]. Significant amount of research has been carried out for new 3D applications. In general, stereoscopic images are acquired from two camera sensors. Displaying the images on a 3D monitor, humans can view and perceive 3D. In the previous applications, the stereoscopic images are delivered to viewers without any modification or enhancement. Any similar efforts have not been performed for solving such problem, yet. Based on this, this paper presents a novel method to enhance 3D perception of the stereoscopic images based upon depth map. The overall aim of the proposed method is to enhance the quality of viewing experience of the end users [3]. *2D+D* (depth map) approach is used as the representation format in our approach. Spatial complexity of depth map is one of the key dimensions by which the perceived quality and depth perception of stereoscopic image are adjusted. The experimental results demonstrate that the lower the spatial complexity is, the higher the perceived video quality and depth perception are. In order to support the assertion, human visual fatigue is also examined.

The paper is organized as follows: Overall approach is introduced in Section 2. Section 3 presents the algorithm of dividing the depth into depth planes utilizing the

spatial complexity of the depth maps. The depth map adjustment algorithm is presented in Section 4. Experimental results are described in Section 5. Finally, Section 6 concludes the paper.

2   Overview of Proposed Method

Fig. 1 shows the overall approach of the proposed method. Given an input depth map, its histogram is analyzed for separating a depth map into multiple depth planes. The spatial complexity is examined for the depth planes. Then the depth planes undergo the adjustment for the variation of 3D perception. Combining the depth planes, a new depth map is made. Finally, a stereoscopic image is generated by 2D+Depth method.
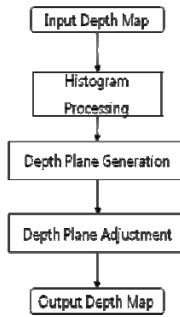

Fig. 1. Block diagram of the proposed method

2.1   Spatial Complexity

Spatial complexity of a depth map is measured by calculating a standard deviation of pixel depth values. The reason behind using the standard deviation for the measurement of spatial complexity is that it is the measure of the dispersion or variability of a set of values around the mean of that set [4]. Thus, if the depth map has high spatial complexity, the standard deviation of the pixel depth values is expected to be high. The pixels in the depth map determine the distance of the associated color image pixel to the viewer. They take grey values ranging from 0 to 255. 0 represents the furthest away pixel from the viewer, while 255 corresponds to the closest pixel to the viewer in a 3D scene.

Given an MxN depth map, the mean pixel depth is computed by

$$\mu_D = \frac{1}{NM} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} D(m,n) \tag{1}$$

Subsequently, the standard deviation is defined by

$$\sigma_D = \sqrt{\frac{1}{NM} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} [D(m,n) - \mu_D]^2} \tag{2}$$

2.2 Histogram Analysis

Depth map histogram $H(i)$ provides the frequency of the depth value i in the depth map D, and is defined as follows:

$$H(i) = \frac{1}{NM} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \delta[i, D(m,n)] \tag{3}$$

where

$$\delta(a,b) = \begin{cases} 1, \text{if } a = b \\ 0, \text{otherwise} \end{cases}$$

As well, a cumulative histogram $C(i)$ of a histogram $H(i)$ is defined as

$$C(i) = \sum_{h=0}^{i} H(h) \tag{4}$$

where C(255) = NM for a 256-level luminance image

2.3 Depth Plane Generation

The histogram separation using Gaussian mixture model has been studied in some applications [5]. This method might work for a couple of objects and background. On the contrary, for images containing multiple objects and background, this method may not work well. Therefore, we use a simple, but efficient method utilizing the cumulative histogram. The following condition is used.

$$| C(i-k) - C(i+k) | < T \tag{5}$$

where T is a threshold value and k is a user-defined parameter.

i values satisfying the above condition are chosen as depth thresholds separating a depth map. Suppose that (L+1) depth thresholds (e.g., i_0, i_1, ... , i_L) are acquired. L depth planes are then generated. Then the range of the lth depth plane is defined as

$$R_D^l = [D_{\min}^l, D_{\max}^l] = [i_{l-1}, i_l], l \in \{1,...,L\} \tag{6}$$

Fig. 2 shows an example of depth thresholds with which a depth map is divided into five depth planes. The red bar indicates the depth thresholds.

Fig. 2. The separation of a depth map into depth planes by depth thresholds (colored in red)

The mean and standard deviation of pixel depth values of each depth plane are computed by

$$\mu^l = \frac{\sum\limits_{j \in R_D^l} D(j)}{No.\,of\ pixels\ in\ R_D^l} \quad and \quad \sigma^l = \sqrt{\frac{\sum\limits_{j \in R_D^l} [D(j) - \mu^l]^2}{No.\,of\ pixels\ in\ R_D^l}} \tag{7}$$

3   Depth Map Adjustment

The block diagram of Fig. 3 shows the depth map adjustment algorithm proposed in this paper. Given depth planes, standard deviation representing the spatial complexity is computed for each depth plane. The source standard deviation σ_S is the sum of depth plane standard deviations. If a target standard deviation σ_T is determined ($\sigma_T < \sigma_S$), the depth range of depth planes are reduced until σ_S is less than σ_T. As a result, the distance between neighboring depth planes are widened and the 3D depth between them becomes stronger. Finally, a stereoscopic image can be generated from 2D+Depth approach.
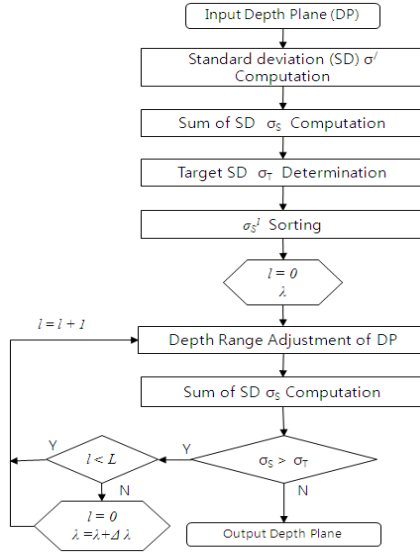

Fig. 3. The block diagram of depth map adjustment method

For each depth plane, the standard deviation σ^l is computed using Eq. (2). Then σ_S is the sum of L depth plane standard deviations.

$$\sigma_S = \sum_{l=1}^{L} \sigma^l \tag{8}$$

To reduce the spatial complexity, we define a target standard deviation σ_T as follows;

$$\sigma_T = \tau \cdot \sigma_S \qquad (9)$$

where τ is a user-defined parameter at [0, 1]. In the experiments, τ is set to be 0.9, 0.8, 0.7, and 0.6.

Until σ_T is achieved, the depth range of depth planes are reduced. The depth planes are sorted according to its standard deviation. The reduction of depth range starts from a depth plane with the greatest standard deviation with a reduction ratio λ. The following equation explains how the range of a depth plane is reduced. For a range [D_{min}^l, D_{max}^l], its depth range is adjusted into [E_{min}^l, E_{max}^l] as follows:

$$E_{min}^l = (1+\lambda)D_{min}^l \quad E_{max}^l = (1-\lambda)D_{max}^l \qquad (10)$$

The depth plane adjustment algorithm is implemented by the following iterative method:

Given L input depth planes,

Step 1: σ^l is computed for each depth plane. Subsequently, σ_S is also computed.

Step 2: Target standard deviation σ_T is set with.

Step 3: We sort the depth planes according to σ^l. The depth adjustment of a depth plane with the greatest standard deviation is processed. $l = 0$ and λ is set to be 0.1 or 0.05.

Step 4: The depth range of l th depth plane is adjusted. As well, a new standard deviation σ_S is also computed.

Step 5: If $\sigma_T < \sigma_S$, examine whether l is less than L. If $l < L$, $l = l + 1$ and go to Step 4. Otherwise l is 0 and increase l by 1 and $\lambda = \lambda + \Delta\lambda$.

Step 6: If $\sigma_T > \sigma_S$, stop and final depth planes are acquired.

4   Experimental Results

The proposed method was performed on various 2D images and depth maps. We illustrate the results for each test image. The first image is MSR *breakdance* image [6] and depth sequences as shown in Fig. 4.

Fig. 4. RGB image and depth map [6]

The depth thresholds are 38, 45, 70, 149, 210, and 216. Five depth planes are shown in Fig. 5.

Fig. 5. Five depth planes of Breakdance

Table 1. Standard deviation of input and output depth planes. (* denotes the range-changed depth planes).

Depth plane	σ^l	Output standard deviation		
		$\tau = 0.9$ $\sigma_T = 998$	$\tau = 0.8$ $\sigma_T = 887$	$\tau = 0.7$ $\sigma_T = 776$
1	43	43	39^*	39^*
2	8	82	68^*	68^*
3	35	359	289^*	188^*
4	41	298^*	176^*	176^*
5	214	214	214^*	214^*
$\sigma_S = \Sigma \sigma^l$	341	996	786	685

When τ is 0.8, at the first iteration, we can not achieve σ_T. So, we increased λ by 0.1. Then in the second iteration, the condition was met at the depth plane 4. A final σ_S is 786. For $\tau = 0.7$, the second iteration was completed with depth plane 3. A final σ_S is 685. Detailed numerical values are found in Table.1. The final depth planes are shown in Fig. 6.

(a)

(b)
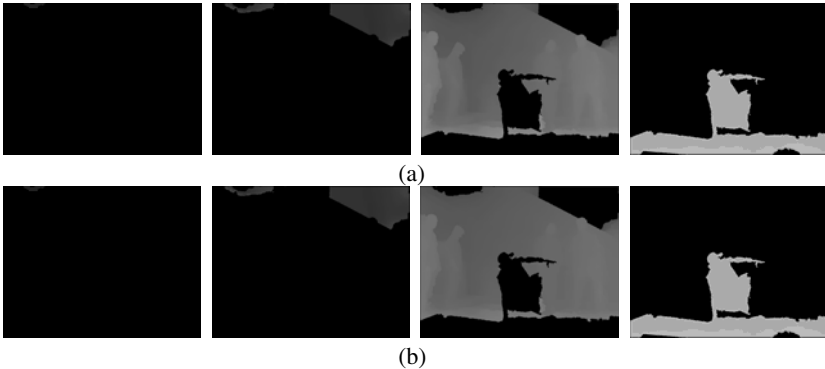
Fig. 6. Depth planes generated according to τ (a) $\tau = 0.8$ and (b) $\tau = 0.7$

Fig. 7 shows the histogram of input and output depth maps. As τ becomes smaller, the distance between depth planes increases. Therefore, depth difference is more apparent.
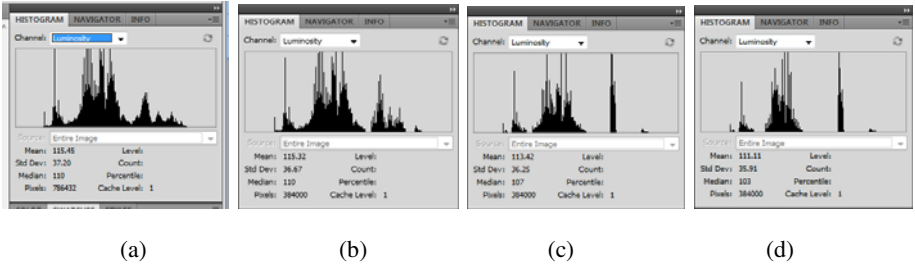

(a)　　　　　　　　(b)　　　　　　　　(c)　　　　　　　　(d)

Fig. 7. The histograms of (a) input depth map and depth maps at (b) τ =0.9, (c) τ =0.8, and (d) τ =0.7

The second test image in Fig. 8 is *Ballet* sequence of MSR [6]. Fig. 9 shows newly adjusted depth planes.
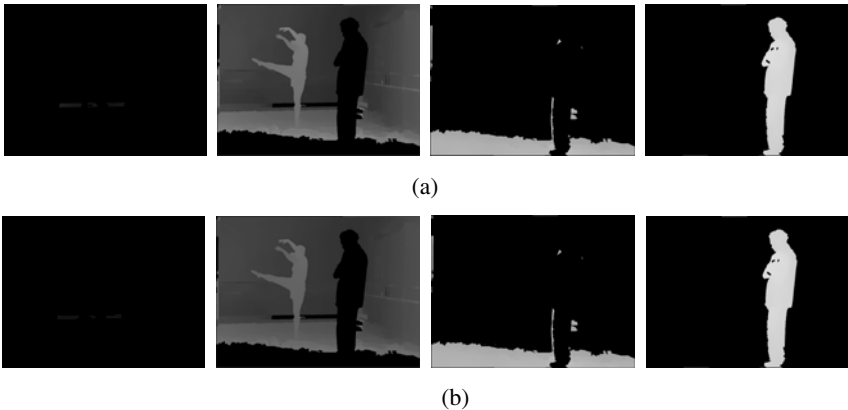

Fig. 8. RGB image and depth map of *Ballet*

(a)

(b)

Fig. 9. Depth planes generated according to τ　(a) τ =0.8 and (b) τ =0.7
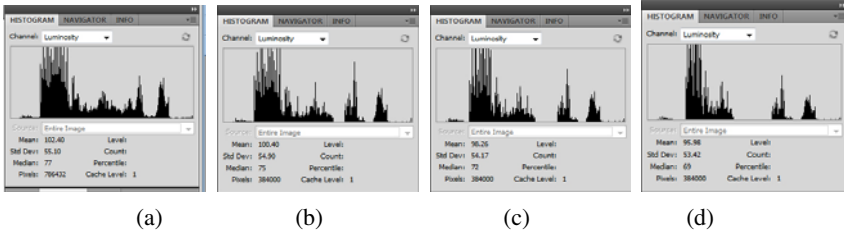
(a)　　　　　(b)　　　　　(c)　　　　　(d)

Fig. 10. The histograms of (a) input depth map and depth maps at τ = (b) 0.9, (c) 0.8, and (d) 0.7

We observed the stereoscopic images with a 3D monitor adopting DQCQS (Double Stimulus Continuous Quality Scale) subjective test [7]. At the first stage, original views were displayed to five participants. Each participant watched the views for 10 seconds and their new views for the same period, and evaluated the effect of the 3D depth. Two test sets were carried out in order to examine the 3D perception improvement. Depth perception was then subjectively judged on a scale of 1 (no improvement), 2 (mild improvement), 3 (average improvement), 4 (good improvement) and 5 (excellent improvement) in terms of 3D perception. Fig. 11 shows two subjective grades with respect to τ as well as σ_S. As τ decreases, we observe that the perceived quality is improved. Furthermore, in order to examine the visual fatigue of the stereoscopic image viewing, we performed subjective visual
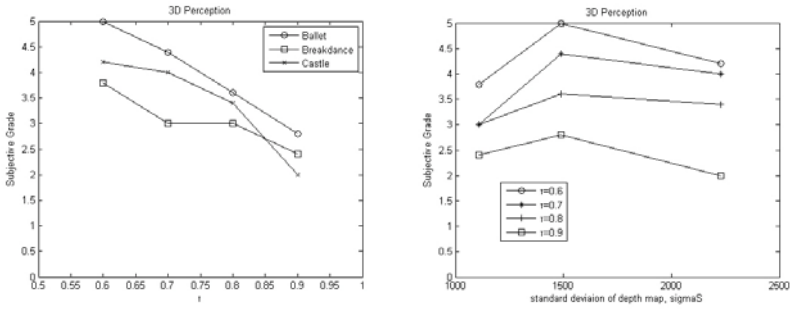

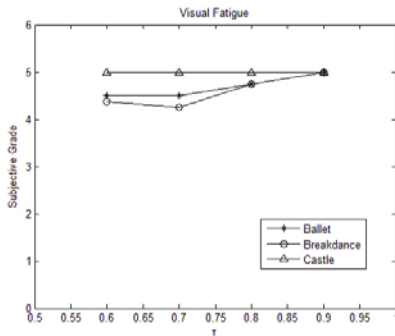**Fig. 11.** Subjective grades with respect to (a) τ and (b) σ_S

Fig. 12. Subjective test for visual fatigue with respect to τ

fatigue test. Visual fatigue was subjectively judged on a scale of 1 (severe fatigue), 2 (fatigue), 3 (mild fatigue), 4 (slight fatigue) and 5 (not at all). As validated in Fig. 12, the overall grade is greater than 4.0, which means that the proposed method makes comfortable stereoscopic images.

5   Conclusion

In this paper, we presented a depth map adjustment method that could provide the improvement of 3D stereoscopic perception. For this, a histogram of a depth map is used for the extraction of multiple depth planes. For spatial complexity, standard deviation of each depth plane is examined. According to the target standard deviation, the depth range of each depth plane is adjusted, thereby making the distance between neighboring depth planes increased. This effect delivers better 3D perception that was validated through subjective tests. Our proposed method is nearly automatic and is expected to provide a technical contribution to 3D video field.

References

1. Meesters, IJsselsteijn, W.A., Seuntiens, P.J.H.: A survey of perceptual evaluations and requirements of three-dimensional TV. IEEE Transactions on Circuits and Systems for Video Technology 14(3), 381–391 (2004)
2. Blonde, L., Doyen, D., Borel, T.: 3D stereo rendering challenges and techniques. In: IEEE 2010 44th Annual Conference on Information Sciences and Systems (CISS), pp. 1–6 (2010)
3. Fehn, C.: Depth-Image-Based Rendering (DIBR), Compression and Transmission for a New Approach on 3D-TV. In: Proc. of the SPIE, CA, U.S.A, vol. 5291, pp. 93–104 (January 2004)
4. Devore, F.: Probability and Statistics for Engineering and the Sciences, Duxbury (1995)
5. Harimi, A., Ahmadyfard, A.: Image Segmentation Using Correlative Histogram Modeled by Gaussian Mixture. In: IEEE Int' Conf. on Image Processing (2009)
6. `http://research.microsoft.com/vision/InteractiveVisualMediaGroup/3DVideoDownload/`, Microsoft Research
7. Lee, E., Heo, H., Park, K.: The comparative measurements of eyestrain caused by 2D and 3D displays. IEEE Trans. on Consumer Electronics 56(3), 1677–1683 (2010)