COMBINING DEPTH INFORMATION AND LOCAL EDGE DETECTION FOR STEREO IMAGE ENHANCEMENT

Walid Hachicha [1] *Azeddine Beghdadi* [1] *and Faouzi Alaya Cheikh* [2]

[1]Université Paris 13, Sorbonne Paris Cité, Institut Galilée, L2TI, France
[2] Gjøvik University College, Media Technology Lab, Norway
walid.hachicha@univ-paris13.fr, azeddine.beghdadi@univ-paris13.fr, faouzi.cheikh@hig.no

ABSTRACT

The recent developments in 3D display technology have opened new horizons and have raised a number of challenges related to the processing and coding of 3D media. Today, stereoscopic image technology is becoming widely used in many fields. The physical limitations of image acquisition systems, however, make stereoscopic technology far from being the most widely accepted solution. Furthermore, the depth/disparity extreme ranges may subject the viewers' eyes to additional strain, causing more discomfort. To address this issue, we propose in this paper to improve stereoscopic image quality by a novel contrast enhancement method that combines local edges and depth information. The contrast is increased locally, at specific depth levels for left and right views. The increase of contrast is controlled based on the depth information, and aims at promoting the nearest objects in the 3D scene. The results obtained from a psychophysical experiment are encouraging and show that the proposed method produces stereo images that are less stressful on the eyes, thus providing more pleasant viewing experience.

Index Terms— Stereo image enhancement, disparity/depth map, local edge detection, human visual system.

1. INTRODUCTION

Stereoscopic imaging technology is becoming increasingly active research field. Despite this renewed interest in stereoscopy, whitnessed over the past 10 years, many problems are still unresolved. Some of these problems are related to the perceptual aspects of stereoscopic imaging. Indeed, viewing 3D scenes may cause eyestrain inducing visual discomfort and headache. Therefore, there is strong interest in solutions allowing the rendering of the 3D visual content giving a pleasant and comfortable viewing experiences. Image enhancement is one of the most promising image processing techniques that could provide such solutions. The enhanced images could be more suitable than the original one for a specific task. In the context of 3D images viewing, enhancing stereoscopic images aims at minimizing or preventing discomfort caused by 3D content and further increasing depth perception. Several enhancement approaches have been proposed for 2D images [1–7], classical 2D state-of-art methods can be roughly classified into two categories: global and local approaches. Global methods are mainly based on intensity mapping using global information such as intensity histogram [1, 2]. Local methods allow the user to modify the contrast based on the information in the vicinity of each pixel. Examples of the second category are contrast enhancement based on edge detection filters [3] and adaptive schemes of unsharp masking for image contrast enhancement [4]. Local contrast enhancement methods proposed in frequency domain, [5, 6]. Furthermore, color vision models such as the Retinex model were also applied for locally amplifying contrast [7]. The applicability to stereo images is limited however, due to their insensitivity to depth information.

Recently, several algorithms have been proposed for stereo image enhancement [8–11]. In [8], authors propose to enhance the depth map using unsharp masking technique by darkening the background of the objects. Jung *et al.* [9] enhance the two views of the stereo images by unsharp masking technique and post-process the images with additional constraints to suppress the changes above the binocular just noticeable difference (BJND) [10]. In [11] the unsharp masking enhancement algorithm is controlled by the disparity.

In this paper we propose to extend the local contrast enhancement algorithm based on edge detection to the case of stereo images, by improving the contrast locally for each object in the 3D scene located at different depth level. In this approach the contrast enhancement amount is modulated by the object position in the 3D scene.

The remainder of this paper is structured as follows. In section 2 we describe the algorithm for contrast enhancement based on local edge detection. We introduce our proposed method based on depth information in section 3. Experimental results that prove the effectiveness of our approach are presented and discussed in section 4. Finally conclusions are drawn in section 5.

2. CONTRAST ENHANCEMENT BASED ON EDGINESS INFORMATION

In this section we give an overview of the contrast enhancement based on local edge detection (CELED) [3], which is the baseline of our work.

CELED is one of the well-known 2D methods in the literature. The authors in [3] proposed a local contrast measurement based on object edge detection. Given $I_{i,j}$ the intensity of the pixel at coordinate (i, j), the local contrast is defined as follows :

$$C_{i,j} = \frac{|I_{i,j} - E_{i,j}|}{I_{i,j} + E_{i,j}}, \quad (1)$$

where $E_{i,j}$ is the mean edge gray level [3] computed by the average value of the pixel gray levels weighted by the edge values $\delta_{m,n}$ over the $(2n+1)\times(2n+1)$ window $w_{i,j}$ centered at the pixel (i, j) and is computed as follows :

$$E_{i,j} = \frac{\sum\limits_{(m,n)\in w_{i,j}} I_{m,n}.\Phi(\delta_{m,n})}{\sum\limits_{(m,n)\in w_{i,j}} \Phi(\delta_{m,n})}, \quad (2)$$

where Φ is an increasing function and $\delta_{m,n}$ the edge value can be estimated by the gradient or Sobel operator output. The image contrast is increased simply by applying a function $\mathbf{f}$ to the local contrast $\mathbf{C}_{i,j}$ to obtain a new contrast $C'_{i,j}$. The function $\mathbf{f}$ satisfies the following conditions:

$$\mathbf{f} : [0, 1] \longrightarrow [0, 1]$$

$$C_{i,j} \longmapsto f(C_{i,j}) = C'_{i,j} \geq C_{i,j}. \quad (3)$$

The resulting intensity $I'_{i,j}$ is deducted as follows :

$$I'_{i,j} = \begin{cases} E_{i,j}.\frac{1-C'_{i,j}}{1+C'_{i,j}} & \text{if } I_{i,j} \leq E_{i,j}, \\ E_{i,j}.\frac{1+C'_{i,j}}{1-C'_{i,j}} & \text{otherwise.} \end{cases} \quad (4)$$

Contrast enhancement based on edge detection algorithm is well adapted to HVS (human visual system). Indeed, it is known that the HVS perception mechanism is very sensitive to edges. Additionally, this algorithm is robust against noise [3]. It improves the gray level distribution and makes easier the extraction of the relevant information. This approach is very useful to discriminate objects according to their boundaries. Moreover, edges become sharper when considering only small windows, otherwise unpleasant effects may appear.

3. PROPOSED DEPTH BASED ALGORITHM

The proposed algorithm aims to exploit both monocular cues, i.e boundaries, contrast, and binocular cue, i.e the disparity between points of the same object of 3D scene viewed from slightly different angles. Combining the local edge detection based algorithm with depth information (depth and disparity are inversely proportional) yields a more perceptually pleasing 3D scene causing less visual discomfort.

The main idea is to increase the contrast over objects located at different depth levels. Note that the sharpness depends also on the object position in the 3D scene. The closer are the objects the sharper they will appear.

Figure1 illustrates the proposed method, which consists of four steps: 1) Disparity estimation, 2) Region growing segmentation, 3) Hole filling and 4) Stereo image processing. The depth information is measured through the disparity between left and right images. Assuming that the disparity results from a horizontal shift, i.e the stereo pair images are taken using a parallel camera model. The relation between two similar pixel in the stereo pair is expressed as :

$$I_{left}(i, j) = I_{right}(i - d_{i,j}, j), \quad (5)$$

where $d_{i,j}$ is the disparity value of the pixel (i, j) in the left view. In the literature, many algorithms have been developed for disparity map estimation, reader can refer to [12] for more details.

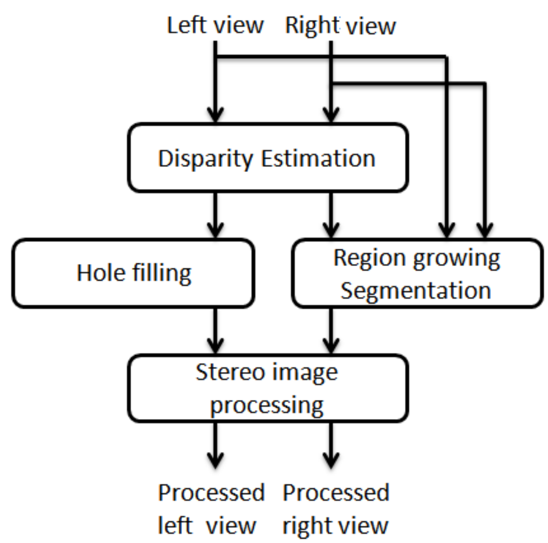

Fig. 1: Flowchart of the proposed stereo image enhancement.

We define an object in the 3D scene as a set of related pixels in the stereo pair images having the same depth. Given a disparity map, we start by detecting objects in the left view. We perform a region growing based segmentation using disparity information. Occluded areas are identified by zero in the disparity map and considered as an object. The objects in the right view are constructed by shifting pixels of objects in the left view using equation 5. Then, for each object $\mathbf{A}$, we compute the mean edge value $\mathbf{E_A}$ given by equation 2. $\mathbf{E_A}$ is identical to $E_{i,j}$ in equation 2 but instead of considering a window of size $(2n+1)\times(2n+1)$ centred at the pixel (i, j), we compute it over the object $\mathbf{A}$. Indeed, making this technique independent of a static window may reduce the appearance of the halo effect caused by the 2D CELED technique. Figure (2) illustrates this phenomenon at object boundaries.

Fig. 2: Midd2 left view [13] enhanced by CELED [3] using 15×15 window size.

An experimental study in [11] shows that the HVS is more sensitive to closer objects in the scene. Based on this observation, we choose a function which decreases with distance, i.e. closer object are more visible. The simplest way is to choose this function as follows :

$$C'_{i,j} = C_{i,j}^{1-\frac{d_{i,j}}{d_{max}}(1-\lambda)}, \tag{6}$$

where $\lambda \in [0, 1]$ is the amount of contrast increase, $d_{i,j}$ is the disparity value of the current pixel and d_{max} is the maximum value of the disparity map. Lowest λ corresponds to the highest contrast. It could be noticed that, any other function satisfying the conditions expressed through equation 3 can be used.

The main problem in stereo matching is the object occlusion and unknown regions after pixel correspondence step. The missing information is identified by zero in the disparity map. To use properly the increasing contrast function, we have to attribute to each pixel a valid disparity value. Gaussian filter based method is a common solution for Hole-Filling in Depth-Image-Based- Rendering (DIBR) [14]. Sometimes hole-size in disparity map is so large that a large Gaussian filter is necessary to complete the missing information. At the same time, this solution involves the loss of the disparity value because of the blur introduced. In our work, hole filling is performed by assigning to each pixel of the hole the average of disparity value surrounding the hole. Figure 3 shows disparity map of the 'Cones' image before and after the proposed hole filling. Once the contrast is raised, the resulting intensity is computed by the inverse function of $C_{i,j}$ given in equation (4).

4. EXPERIMENTS AND RESULTS

In this section we report results based on a widely used stereo dataset [13]. Ten stereo images are chosen with different

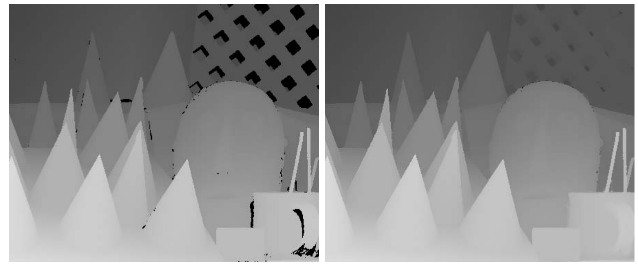

Fig. 3: Cones image disparity map before and after hole filling.

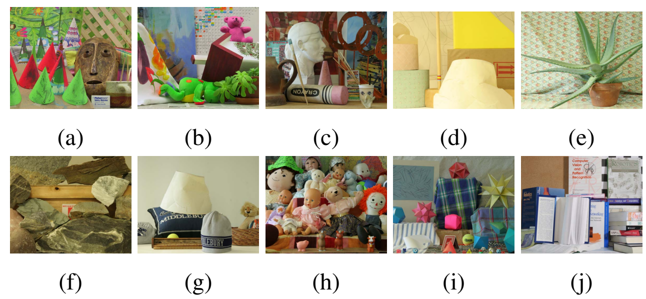

Fig. 4: Left views of the stereo pair dataset used in the subjective test collected from [13], (a) Cones, (b) Teddy, (c) Art, (d) Lampshade2, (e) Aloe, (f) Rock1, (g) Midd2, (h) Dolls, (i) Moebius and (j) Books.

complex geometry, occlusion and having the same resolution 463 x 370 except 'Cones' and 'Teddy' that are in size 448 x 374, see figure 4. In our experiment we take profit of ground truth disparity acquired using techniques that employ structured lighting [15]. Obviously, our choice does not prevent the use of any algorithms to estimate the disparity map, however efficient ones are strongly recommended. The middlebury benchmark website [13] gives a rank order of existing methods. A test bed for quantitative evaluation of stereo algorithms is given in [12].

The objective of the experiment is to emphasize the impact of the depth information on the enhancement process. We consider the 2D CELED method as baseline for comparison.

Original RGB stereo pairs were converted to Lab color space with L being the luminance and a, b the corresponding red-green and blue-yellow channels, respectively. Note that only luminance channel information is processed here.

In general, image enhancement needs to be evaluated by subjective quality assessment. The definition of specific test setups for subjective test experiments is required. Methods have been proposed for 2D quality, especially the double-stimulus continuous quality scale (DSCQS) recommended in ITU-R BT.500 is widely used [16]. The stereo pairs are displayed using ACER GD245HQ with active liquid crystal shutter glasses, 18 observers, mostly males considered as experts, with an average age of 31 years and have normal or correct-to-normal visual acuity. Each subject was individually briefed about the experiment, and a demonstration of the experimental procedure was given. Subjects were asked to rank the processed stereo images by 2D CELED and our proposed algorithm, viewing the original image as reference. Numerous criteria are taken into account for the ranking procedure, including quality of the stereo image, perceived depth and visual strain.

Each enhanced stereo image pair is juxtaposed with reference (original stereo pair) separated by a vertical line. Juxtaposed images are displayed for ten seconds in random order. A uniform grey tone screen is displayed, in between the two images to be evaluated, for two seconds and five seconds are left for the observer to vote.

Figure 5 shows the subjective test results described above. It represents for each stereo pair the frequency of the first rank for both 2D CELED and the proposed method. It is shown that the proposed enhancement technique outperforms the 2D CELED method for all images. The proposed method produces more pleasant stereo images and comfortable to visualise as reported by the majority of observers, more than 90% on average. In fact, the observer focuses on the closer object more than the farthest ones. Consequently, they provide less effort than focusing on the whole 3D scene. Taking into account the depth information makes the enhanced stereo images more realistic and increases the depth perception.

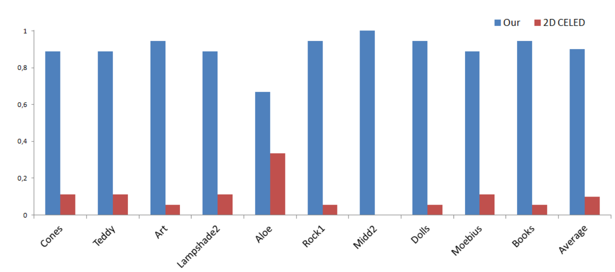

Fig. 5: Ranking results for both the 2D CELED method [3] and our proposed method.

Figure 6 illustrates the left view of the Cones, Books and Rock1 stereo pairs [13] enhanced by the proposed algorithm and the 2D CELED method, $\lambda = 0.7$ for the two methods and we consider $\frac{d}{d_{max}} = 1$ for the 2D CELED and the window size is equal to 15×15. We can even notice that for one view of the stereo pairs the halo effect does not appear in these views processed by our method unlike classical 2D methods. In addition our method generates views with less distortion when compared to the original, closer objects are sharper whereas farther away objects appear increasingly blurred. While in the 2D method, the contrast increases uniformly over the whole view.

5. CONCLUSIONS

We introduced a novel stereo image contrast enhancement approach which combines the sensibility of the HVS to edge information and position of objects in the 3D scene. The experiments carried out on a widely used stereo image dataset show the effectiveness of the proposed method in using the depth information in stereo enhancement process. Our method produces pleasant 3D scenes increases the perception of depth and causes less eye strain. In this work we focused only on edge and depth information. The introduction of visual saliency, some local features and geometrical information in the enhancement process will be considered in a future work.

6. REFERENCES

[1] Robert Hummel, "Image enhancement by histogram transformation," *Computer Graphics and Image Processing*, vol. 6, no. 2, pp. 184 – 195, 1977.

[2] J.A. Stark, "Adaptive image contrast enhancement using generalizations of histogram equalization," *IEEE Transactions on Image Processing*, vol. 9, pp. 889–896, 2000.

[3] A. Beghdadi and A. Le Négrate, "Contrast enhancement technique based on local detection of edges," *Computer Vision, Graphics, and Image Processing*, vol. 46, no. 2, pp. 162–174, 1989.

[4] A. Polesel, G. Ramponi, and V.J. Mathews, "Adaptive unsharp masking for contrast enhancement," in *Proceedings of International Conference on Image Processing*, 1997, pp. 267–271.

[5] J-L Starck, F. Murtagh, E.J. Candes, and D.L. Donoho, "Gray and color image contrast enhancement by the curvelet transform," *IEEE Transactions on Image Processing*, vol. 12, 2003.

[6] D. Cherifi, A. Beghdadi, and A.H. Belbachir, "Color contrast enhancement method using steerable pyramid transform.," *International Journal of Signal, Image and Video Processing*, vol. 4, no. 2, pp. 247–262, 2010.

[7] L. Meylan and S. Süsstrunk, "Bio-inspired color image enhancement," in *Proceedings of SPIE: Human Vision and Electronic Imaging*, 2004, pp. 46–56.

[8] T. Luft, C. Colditz, and O. Deussen, "Image enhancement by unsharp masking the depth buffer," *ACM Transactions on Graphics*, vol. 25, no. 3, pp. 1206–1213, 2006.

[9] S.W. Jung, J.Y. Jeong, and S.J. Ko, "Sharpness enhancement of stereo images using binocular just-noticeable-difference," *IEEE Transactions on Image Processing*, 2011.

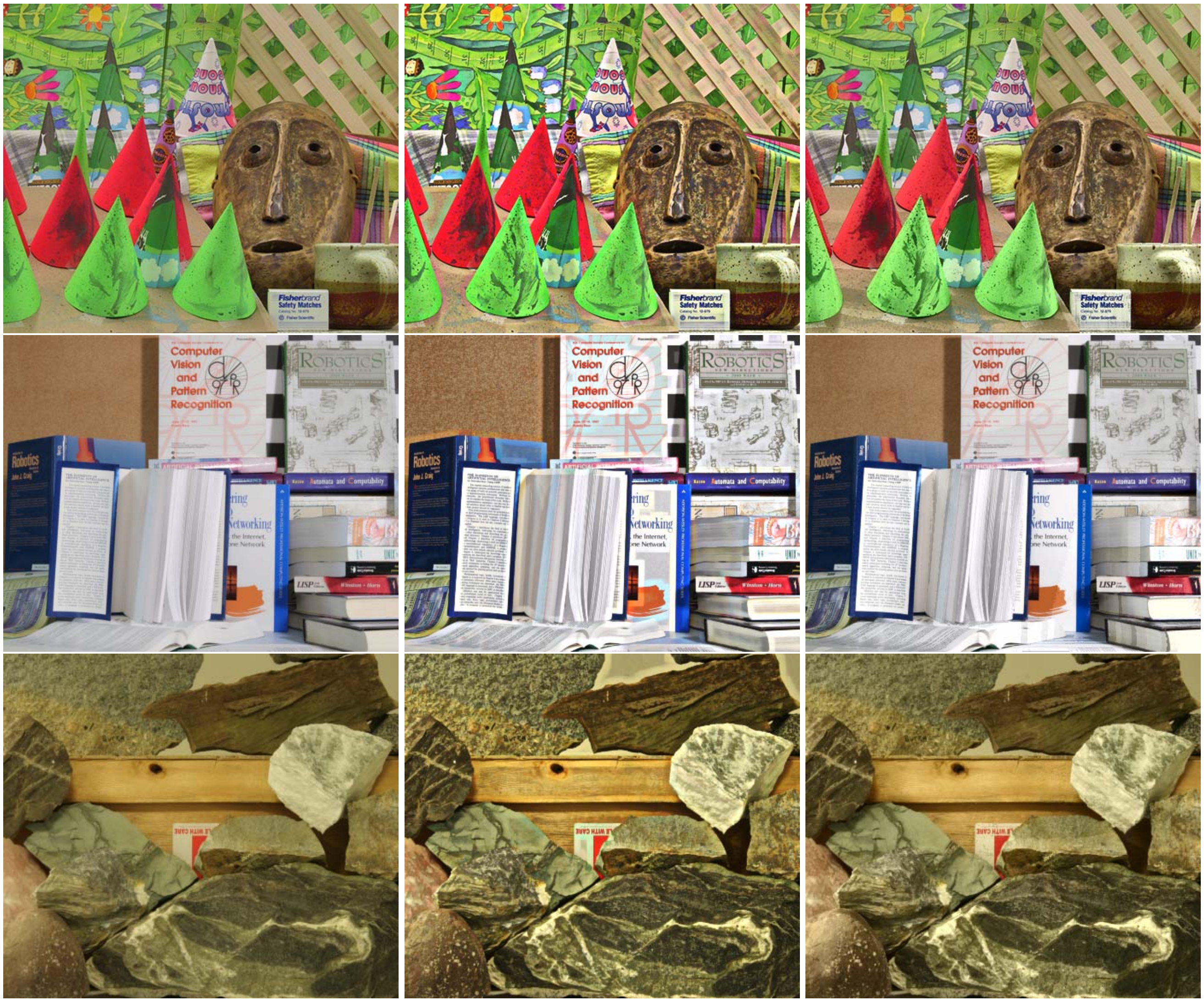

Fig. 6: Column 1: Original left view, column 2: processed left view by CELED 2D, column 3: processed left view by proposed algorithm.

[10] Y. Zhao, Z. Chen, C. Zhu, Y.P. Tan, and L. Yu, "Binocular just-noticeable-difference model for stereoscopic images," *IEEE Signal Process*, vol. 18, NO.1, pp. 19–22, 2011.

[11] M.M. Subedar and L.J. Karam, "Increased depth perception with sharpness enhancement for stereo video," *Displays*, vol. 7524, no. Figure 1, pp. 75241B–75241B–8, 2010.

[12] D. Scharstein and R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms," *International Journal of Computer Vision*, vol. 47, pp. 7–42, 2001.

[13] D. Scharstein and R. Szeliski, "Middleburry stereo vision website," *http://vision.middlebury.edu/stereo*.

[14] L. Zhang, W.J. Tam, and D. Wang, "Stereoscopic image generation based on depth images," in *Proceedings of International Conference on Image Processing*, 2004, pp. 2993–2996.

[15] D. Scharstein and R. Szeliski, "High-accuracy stereo depth maps using structured light," in *Proceedings of the 2003 IEEE computer society conference on Computer vision and pattern recognition*, 2003, pp. 195–202.

[16] "Methodology for the subjective assessment of the quality of television pictures," *ITU-R BT . 500-11*, 2002.