EVALUATION OF LDR, TONE MAPPED AND HDR STEREO MATCHING USING COST-VOLUME FILTERING APPROACH

Tara Akhavan, Hyunjin Yoo, Margrit Gelautz

Institute of Software Technology and Interactive Systems
Vienna University of Technology, Austria

ABSTRACT

We present stereo matching solutions based on a fast cost-volume filtering approach for High Dynamic Range (HDR) scenes. Multi-exposed stereo images are captured and used to generate HDR and Tone Mapped (TM) images of the left and right views. We perform stereo matching on conventional, Low Dynamic Range (LDR) images, original HDR, as well as TM images by customizing the matching algorithm for each of them. An evaluation on the disparity maps computed from the different approaches demonstrates that stereo matching on HDR images outperforms conventional LDR stereo matching and TM stereo matching, with the most discriminative disparity maps achieved by using *HDR color information* and *log-luminance gradient values* for matching cost calculation.

Index Terms— Stereo Matching, Low Dynamic Range (LDR), High Dynamic Range (HDR), Tone Mapping (TM).

1. INTRODUCTION

Stereo matching, that is the identification of corresponding points between the left and right image of a stereo pair, has been an active area of computer vision research for many years. There are some well-known challenges in stereo matching such as low-textured regions, repeating patterns, saturated areas and illumination changes between the two stereo views. Capturing a huge range of illumination in High Dynamic Range (HDR) scenes using only 256 values per color channel - as in conventional Low Dynamic Range (LDR) images - typically causes over- and/or under-exposed regions in the image. Many of the available state-of-the-art stereo matching methods perform well in absence of the mentioned regions, however, it is not possible to avoid these challenging areas in real world imaging. A lot of research has tried to find solutions to these problems [1, 2]. In Fig. 1 some of the challenging areas for stereo matching are highlighted on an example from our self-captured stereo data set. Conventional stereo matching methods are unlikely to perform well on an HDR scene as shown on an example in Fig. 1 (d).

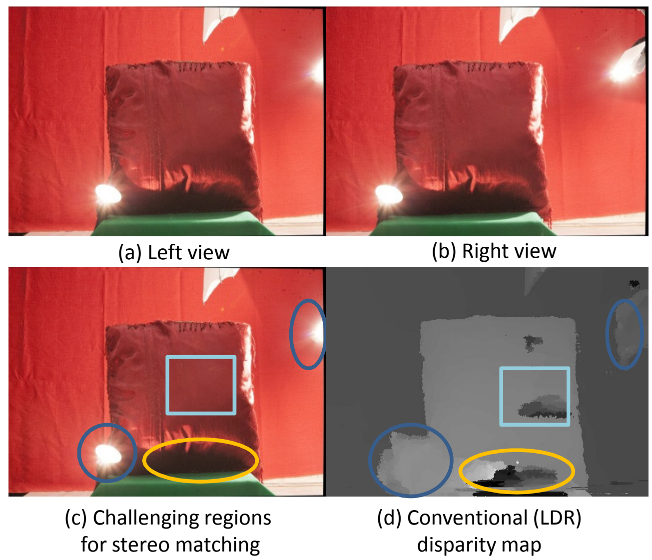

Fig. 1. Example of an HDR scene containing challenging regions for stereo matching. The left and right views are shown in (a) and (b). In (c), regions specified in dark blue, yellow and light blue, respectively, highlight over-exposed, under-exposed and low-textured areas in the left image. The disparity map (d) is calculated using the cost-volume filtering method [6]. Baseline: 75 mm.

There has been a trend toward HDR imaging in recent years [3, 4]. Images that store a depiction of the scene in a range of intensities proportional to the scene radiance are called HDR or radiance maps [5]. Tone Mapping (TM) operators are designed to compress the contrast of HDR images into a limited dynamic range of displays or printers. In this paper we use multi-exposed images of each view to experiment with different stereo matching approaches on three types of imagery (LDR, TM, HDR). Our main goal is to find an appropriate solution to achieve informative disparity maps of HDR real world scenes.

Only a small amount of literature is available on the joint area of HDR imaging and stereo matching, mostly focused on using depth information to generate better HDR images/videos [7–10] or capturing stereoscopic HDR data [11, 12]. In contrast, our goal is to exploit the radiance information in HDR images to obtain disparity maps of higher

Tara Akhavan is funded by the Vienna PhD School of Informatics. Hyunjin Yoo has been supported by the Vienna Science and Technology Fund (WWTF) through project ICT08-019.

quality. Akhavan et al. [13] outlined a theoretical framework for solving the multi-exposed stereo matching problem (see Fig. 2). The authors suggested using HDR images generated for the left and right views and tone mapped stereo pairs as inputs to the stereo matching process. In this paper:

- We implement and extend the Akhavan et al. approach using two different tone mapping methods as well as three different ways of applying stereo matching on HDR stereo pairs.
- We introduce a stereo data set of conventional-tone mapped-high dynamic range (LDR-TM-HDR) images and compare the disparity maps obtained from the three different groups.

In Section 2, we describe our data set and matching approaches, and in Section 3, our stereo matching evaluation performed on the LDR-TM-HDR stereo data set is illustrated and discussed.

2. DATA SET AND STEREO MATCHING APPROACHES

We follow the stereo matching method introduced in [6] as our base stereo matching approach and customize it with our various proposed scenarios. In this section, our self-captured stereo data set as well as our proposed LDR, TM and HDR stereo matching methods are discussed.

2.1. LDR-TM-HDR stereo data set

We generated a stereo data set containing the challenging regions for stereo matching as shown in Fig. 1. Capturing a scene in multiple exposures, when the camera and the scene are both still, is very common [14] and does not require so much effort or time since most of the cameras can be configured for this purpose. Our data set consists of:

- LDR stereo images in 8 different exposures of: 1/15s, 1/30s, 1/60s, 1/125s, 1/250s, 1/500s, 1/1000s and 1/2000s.
- Drago tone mapped stereo images.
- Durand tone mapped stereo images.
- HDR stereo images.
- Log-HDR stereo images.

All of the stereo pairs are rectified to achieve epipolar geometry and simplify the search from two dimensions into one dimension. Fig. 3 shows an HDR image constructed from multi-exposed images. Two samples of the data set are shown in Fig. 4. For the LDR images only the middle exposure is shown. To our knowledge, there is no HDR stereo data set publicly available. Note that the well-known Middlebury stereo data set[1] cannot be used for our experiments since it does not contain HDR scenes.

[1]http://vision.middlebury.edu/stereo/data/

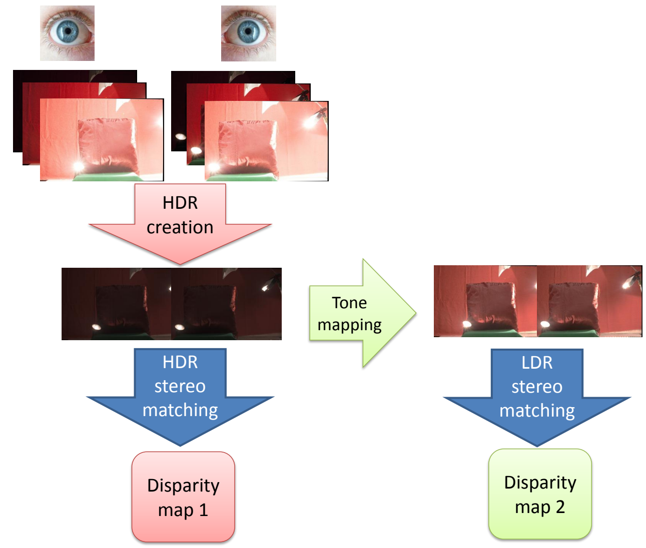

Fig. 2. Two general approaches for solving the multi-exposed stereo matching problem in HDR scenes.

2.2. LDR stereo matching

One could use the conventional stereo matching methods with the conventional LDR stereo input images of the HDR scene. This would result in disparity maps such as the one presented in Fig. 1. We used a state-of-the-art local stereo matching technique based on cost-volume filtering [6]. The matching cost calculation using the color values (Ic_i) and grayscale gradient information ($\bigtriangledown_x I_i$) is formulated in Eq.1 [6].

$$C_{i,d} = \alpha \cdot min[\|Ic_i - Ic'_{i-d}\|, \tau_1] + \\ (1-\alpha) \cdot min[\|\bigtriangledown_x I_i - \bigtriangledown_x I'_{i-d}\|, \tau_2]. \quad (1)$$

The cost-volume entry, $C_{i,d}$, determines how well a pixel i in the left image matches the same pixel in the right image shifted by vector (disparity) d in the x direction. Here, $\bigtriangledown_x$ is the gradient operator in the x direction. For weighting the color and gradient information α is used, and τ values are truncation values. We refer to this approach as *LDR stereo matching* in the rest of the paper.

2.3. TM stereo matching

As presented in Fig. 2, we exploit the information from multi-exposed images of both left and right views for our proposed stereo matching approaches. HDR images of each view are calculated from multi-exposed bracketing following the approach in [14]. Then a tone mapping operator can be applied to compress the HDR image back to the range of 256 values. The tone mapped image is treated like an LDR image and can be used easily as an input to any stereo matching method. In most of the tone mapping evaluation papers, Drago et al. [15] and Durand and Dorsey [16] methods are considered among the most effective operators [17–19]. We applied these two well known tone mappers on our HDR stereo data and

Fig. 3. Sample of an HDR image (dynamic range = 8150:1) computed from 8 exposures. The highest, middle and lowest exposed images are shown in the bottom row.

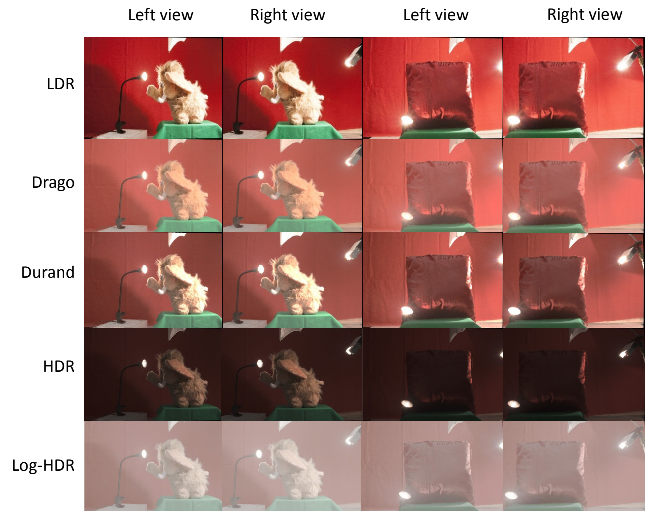

Fig. 4. Samples of LDR mid-exposed, two different tone mapped, original HDR, and log-HDR stereo images. Baseline: 150 mm.

computed the disparity map using the cost values calculated from Eq.1 after appropriate filtering [6]. Tone-mapping operators have been developed for display purposes and the tone-mapped images are meant to contain more visible detail than LDR images using the same number of bits. The Drago et al. tone mapping operator is a global approach while the Durand and Dorsay represents a local method. In the former approach, the contrast is reduced using a logarithmic compression of luminance values. The latter method decomposes the HDR image into a base layer and a detail layer using an edge-preserving bilateral filter and compresses only the base layer contrast.

2.4. HDR stereo matching

Another approach suggested in Fig. 2, is to use the original HDR images in stereo matching. Using HDR images as inputs to stereo matching needs customization in matching algorithms.

2.4.1. HDR approach

To compute the stereo matching costs by using HDR images, we replace the RGB of the LDR image (Ic_i) and the gradient of the intensity ($\bigtriangledown_x I_i$) in Eq. 1 with the RGB values from the HDR image (R_i) and the luminance gradient values ($\bigtriangledown_x Lv_i$), respectively, as shown in Eq. 2.

$$C_{i,d} = \alpha{\cdot}min[\|R_i - R'_{i-d}\|, \tau_1]+ \\ (1-\alpha){\cdot}min[\|\bigtriangledown_x Lv_i - \bigtriangledown_x Lv'_{i-d}\|, \tau_2]. \quad (2)$$

2.4.2. HDR logarithmic approach

Since human vision perception of light and illumination changes could be imitated by a logarithmic function [20], we simply apply the log operator on the stereo HDR images as in Eq. 3.

$$C_{i,d} = \alpha{\cdot}min[\|\log(R_i) - \log(R'_{i-d})\|, \tau_1]+ \\ (1-\alpha){\cdot}min[\|\bigtriangledown_x \log(Lv_i) - \bigtriangledown_x \log(Lv'_{i-d})\|, \tau_2]. \quad (3)$$

2.4.3. HDR RGB and log-luminance approach

As an alternative approach, we found it more effective to keep the original RGB values from the HDR image and apply the log function only on luminance values as formulated in Eq. 4.

$$C_{i,d} = \alpha{\cdot}min[\|R_i - R'_{i-d}\|, \tau_1]+ \\ (1-\alpha){\cdot}min[\|\bigtriangledown_x \log(Lv_i) - \bigtriangledown_x \log(Lv'_{i-d})\|, \tau_2]. \quad (4)$$

3. EXPERIMENTAL RESULTS

In this section, we show, discuss and compare the results of all proposed approaches towards stereo matching in HDR scenes. All disparities are computed for the left view. Since no ground truth disparity map is available at the moment, we use a qualitative evaluation approach by visual comparison of the results. Fig. 5 gives an overview of the results obtained from the different approaches. The disparity maps are shown in gray value encoded representation, with brighter values associated with larger disparities. Our main observations according to Fig. 5 are discussed in more detail in the following:

1. LDR disparities presented in the third row clearly show matching problems in low-textured and under-exposed areas (marked as yellow), differently illuminated regions[2]

[2] Note that a lamp that is not visible in the images is placed on the right side of the scene, which causes more illumination changes in the right border areas of the stereo pairs.

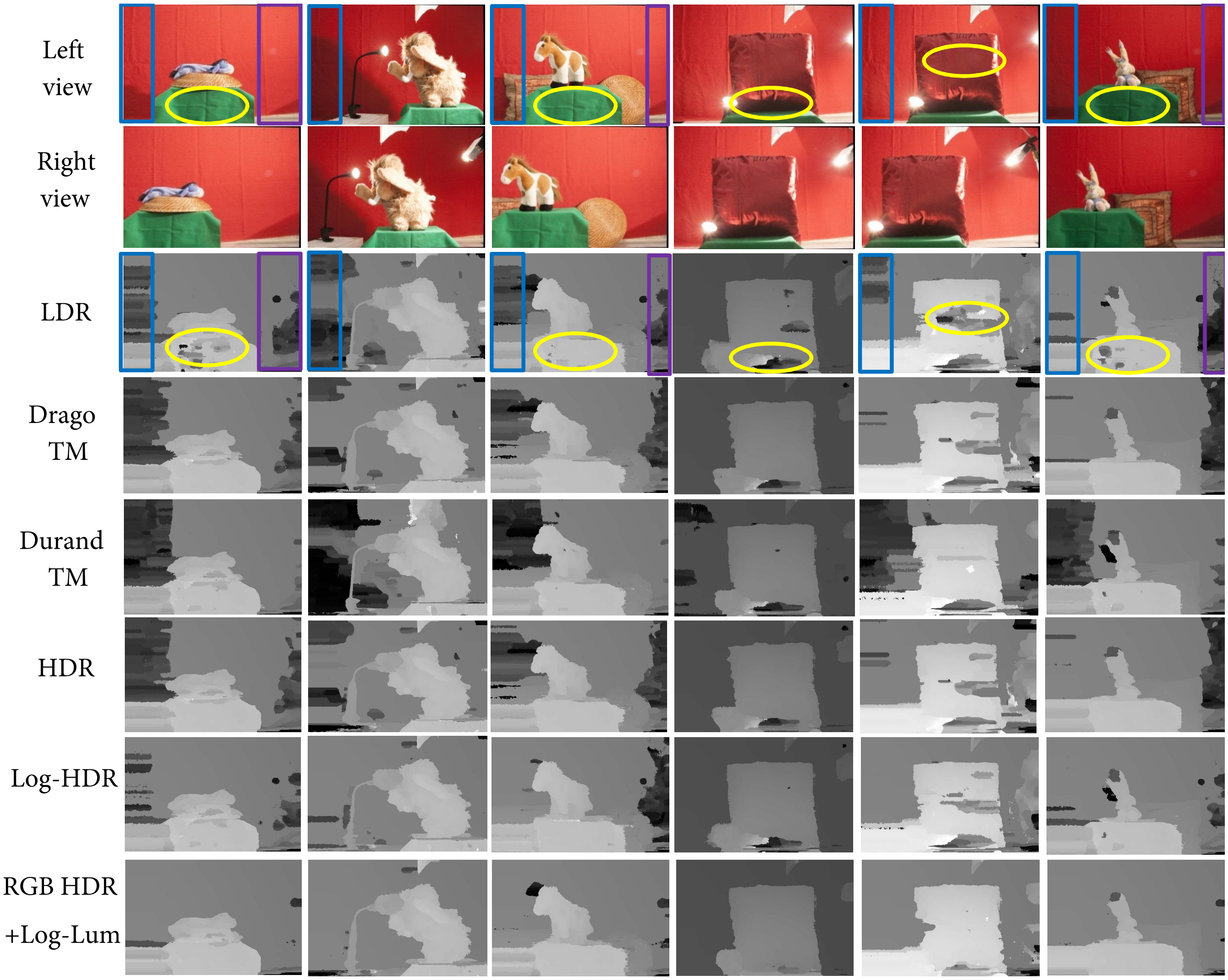

Fig. 5. LDR, TM (two different methods), and HDR (three different methods) disparity map comparison. Some of the low-textured and under-exposed regions are marked as yellow, some areas with more pronounced illumination differences between the stereo pairs are marked as purple, and some filled in areas (border regions) are marked as blue. Baselines: 150 mm, 150 mm, 150 mm, 75 mm, 150 mm, 150 mm.

(marked as purple) and filled in regions (occluded regions that are not visible in both views, filled in during the post processing phase of the stereo matching [6]) of the HDR scenes.

2. Our experiments show that applying stereo matching on tone mapped image pairs (rows 4 and 5) improves the disparity results in the yellow marked areas but not usually in blue and purple marked ones.
3. All three HDR stereo results (rows 6-8) show improvements in the yellow marked areas compared to the LDR results. An improvement with respect to the TM results is also visible in the two last rows for most cases (see areas marked blue and purple). Among the three introduced HDR approaches, using HDR RGB values and log-luminance gradient information for the matching cost calculation, as discussed in 2.4.3, provides the best disparity maps. Their results (shown in the last row) outperform the LDR and TM results for all test cases.
4. Bigger base lines between stereo pairs, specially in HDR scenes with large illumination changes, are more challenging. The two pillow images with the same left view and different right views of different baselines (columns 4 and 5) show how a bigger base line between the image pairs in HDR scenes causes more challenges to matching in all three approaches.

Our experimental results indicate that computing the stereo matching in the HDR space leads to better disparity information. For this purpose, one should consider a customization process needed in the stereo matching algorithms as described in Section 2.4 as well as the need for dealing with a higher range of pixel values (not between 0 and 255 as conventional images). This customization will usually depend on the stereo matching approach. In this paper we suggested three possible ways to customize a specific matching approach, which means HDR stereo matching is not backward compatible with the existing matching algorithm. In contrast, using HDR tone-mapped stereo images is a backward compatible approach with the existing stereo matching techniques. In our experiments, however, the gain in quality obtained from the tone-mapped stereo images was less noticeable than that achieved by stereo matching in HDR space.

4. CONCLUSION

In this paper, we discussed the problem of stereo matching in HDR scenes. We generated our own LDR-TM-HDR stereo data set using two different tone mapping operators and customized the stereo matcher to the HDR images by suitable modification of the cost computation. The experiments showed that (1) using HDR images results in better disparity maps in comparison to conventional LDR inputs and tone mapped ones, and (2) taking into account the original RGB values from the HDR image with the log-luminance gradient values, for computing the cost-volume, achieves the best results among all compared approaches.

REFERENCES

[1] H. Hirschmüller, "Accurate and efficient stereo processing by semi-global matching and mutual information," in *CVPR*, 2005, pp. 807–814.

[2] Q. Yang, C. Engels, and A. Akbarzadeh, "Near real-time stereo for weakly-textured scenes," in *BMVC*, 2008, pp. 72.1–72.10.

[3] Y. Zhang, E. Reinhard, and D. R. Bull, "Perceptually lossless High Dynamic Range image compression with JPEG 2000," in *ICIP*, 2012, pp. 1057–1060.

[4] E. Selmanovic, K. Debattista, T. Bashford-Rogers, and A. Chalmers, "Backwards compatible JPEG stereoscopic high dynamic range imaging," in *TPCG*, 2012, pp. 1–8.

[5] E. Reinhard, G. Ward, S. Pattanaik, P. Debevec, and W. Heidrich, *High Dynamic Range Imaging - Acquisition, Display, and Image-based Lighting (2. ed.)*, Academic Press, 2010.

[6] A. Hosni, C. Rhemann, M. Bleyer, C. Rother, and M. Gelautz, "Fast cost-volume filtering for visual correspondence and beyond," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 2, pp. 504–511, 2013.

[7] A. Troccoli, S. B. Kang, and S. M. Seitz, "Multi-view multi-exposure stereo," in *3DPVT*, 2006, pp. 861–868.

[8] N. Sun, H. Mansour, and R. K. Ward, "HDR image construction from multi-exposed stereo LDR images," in *ICIP*, 2010, pp. 2973–2976.

[9] D. Rüfenacht, "Stereoscopic high dynamic range video," M.S. thesis, EPFL, Lausanne, Switzerland, 2011.

[10] J. Bonnard, C. Loscos, G. Valette, J. Nourrit, and L. Lucas, "High-dynamic range video acquisition with a multiview camera," in *SPIE*, 2012, vol. 8436, pp. 84360A–1–84360A–11.

[11] E. Selmanović, K. Debattista, T. Bashford-Rogers, and A. Chalmers, "Generating stereoscopic HDR images using HDR-LDR image pairs," *ACM Trans. Appl. Percept.*, vol. 10, no. 1, pp. 3:1–3:18, 2013.

[12] E. Selmanović, K. Debattista, T. Bashford-Rogers, and A. Chalmers, "Enabling stereoscopic high dynamic range video," *Signal Processing: Image Communication*, vol. 29, pp. 216–228, 2014.

[13] T. Akhavan, H. Yoo, and M. Gelautz, "A framework for HDR stereo matching using multi-exposed images," in *HDRi2013 - First International Conference and SME Workshop on HDR imaging*, 2013.

[14] P. Debevec and J. Malik, "Recovering high dynamic range radiance maps from photographs," in *ACM SIGGRAPH*, 1997, pp. 369–378.

[15] F. Drago, K. Myszkowski, T. Annen, and N. Chiba, "Adaptive logarithmic mapping for displaying high contrast scenes," *Computer Graphics Forum*, vol. 22, pp. 419–426, 2003.

[16] F. Durand and J. Dorsey, "Fast bilateral filtering for the display of High-dynamic-range images," *ACM Trans. Graph.*, vol. 21, pp. 257–266, 2002.

[17] A. Yoshida, V. Blanz, K. Myszkowski, and H. Seidel, "Perceptual evaluation of tone mapping operators with real-world scenes," in *Human Vision and Electronic Imaging X, SPIE*, 2005.

[18] P. Ledda, A. Chalmers, T. Troscianko, and H. Seetzen, "Evaluation of tone mapping operators using a high dynamic range display," in *ACM SIGGRAPH*, 2005, pp. 640–648.

[19] M. Čadík, M. Wimmer, L. Neumann, and A. Artusi, "Image attributes and quality for evaluation of tone mapping operators," in *14th Pacific Conference on Computer Graphics and Applications*, 2006, pp. 35–44.

[20] J.E. Roeckelein, *Dictionary of Theories, Laws, and Concepts in Psychology*, Greenwood Press, 1998.