YDA Görüntüler İçin Gölgeleme Giderme Algoritmalarının Karşılaştırılması
Evaluating Deghosting Algorithms for HDR Images

Okan Tarhan Tursun, Ahmet Oğuz Akyüz
Bilgisayar Mühendisliği Bölümü
Orta Doğu Teknik Üniversitesi
Ankara, Türkiye
{tarhan,akyuz}@ceng.metu.edu.tr

Aykut Erdem, Erkut Erdem
Bilgisayar Mühendisliği Bölümü
Hacettepe Üniversitesi
Ankara, Türkiye
{aykut,erkut}@cs.hacettepe.edu.tr

***Özetçe* —Gerçek hayat koşullarında görüntülenen sahnelerdeki ışınım yeğinliği büyük değişkenlik gösterebilir. Işınım yeğinliğindeki değişkenliği doğru bir şekilde yakalamak için Yüksek Dinamik Aralıklı (YDA) görüntüleme teknikleri kullanılmaktadır. Bu tekniklerin bir kısmı aynı sahnenin farklı pozlama süreleriyle çekilmiş Düşük Dinamik Aralıklı (DDA) birden fazla görüntüsünü girdi olarak alarak bir YDA görüntü üretmeyi temel alır. Görüntüleme ve birleştirilme işlemi sırasında DDA görüntülerin arasında kamera ve sahnedeki objelerin hareketinden doğabilecek farklılıkların, kullanılan YDA oluşturma yöntemi tarafından doğru bir şekilde çözümlenmesi gerekir. Bu çalışmada bu probleme yönelik olarak son yıllarda önerilen iki yaklaşım, değişik zorlukta resim dizilerinden oluşacak şekilde ürettiğimiz bir veri kümesinde, nesnel ve öznel değerlendirme ölçütleri temel alınarak sınanmış ve sonuçları paylaşılmıştır.**

Anahtar Kelimeler—YDA görüntüleme, gölgeleme giderme, nesnel değerlendirme, öznel değerlendirme.

***Abstract*—The real world encompasses a high range of luminances. In order to capture and represent this range correctly, High Dynamic Range (HDR) imaging techniques are introduced. Some of these techniques are based on constructing an HDR image from several Low Dynamic Range (LDR) images with different exposures. In the capture and reconstruction phases, the HDR reproduction techniques must resolve the differences between the input LDR images due to camera and object movement. In this study, two recent approaches addressing this issue are compared using a novel dataset comprised of image sequences with varying complexity. The results are evaluated by using both objective and subjective measures.**

Keywords—HDR imaging, deghosting, objective evaluation, subjective evaluation.

I. INTRODUCTION

In both amateur and the professional imaging applications, there may be a large variance in the scene irradiance. Capabilities of the traditional LDR imaging equipments have a limitation on the difference of minimum and maximum scene irradiance which may be captured correctly. When this difference falls out of the capability of the used sensor, overexposed and underexposed pixels are observed in the image. Presence of a strong light source such as the sun, a lamp, flame and

(a)

(b)

Figure 1: Ghosting problem: (a) Camera with motion, (b) Objects with motion.

the presence of objects in a dimly illuminated regions such as those under shadow or in dark regions may be an example to such scenes [1].

Mitsunaga and Nayar [2], Guo and Sonkusale [3], Tocci et al. [4] proposed hardware specially designed for this problem. In addition, some manufacturers introduced commercial imaging equipments with dynamic ranges larger than the traditional equipments [5]. However, these products are both do not seem to replace the traditional imaging equipments in the near future since they are costly and they also provide a lower dynamic range compared to the exposure bracketing. Therefore, in order to produce an HDR image, taking a sequence of LDR images with different exposures became a popular method.

Many approaches proposed for this HDR construction technique are based on the assumption that the input LDR images have a pixelwise overlap and correspondence [6]–[10]. However, in the real world when the objects in the scene or the acquisition equipment are not stationary, this assumption is not valid, leading to ghosting artifacts in the output HDR image (Figure 1).

There are different studies conducted to eliminate the ghosting artifacts. Zimmer et al. [11] proposes an energy-based approach to find the optical flow between the input LDR images and to align input images. In order to overcome the effects of the different exposure settings, they used the vertical and horizontal image gradients which are relatively less sensitive to the changes in the exposure. As a result of the alignment using the dense correspondence map obtained by energy minimization, HDR images with superresolution are produced.

In their study, Hu et al. [12] picks the image with the smallest number of under and overexposed pixels as the reference image from the input LDR image set and estimate a color transfer function between the reference image and remaining images. Their proposed approach finds a pixelwise correspondance based on the method of HaCohen et al. [13]. The under and overexposed pixels in the reference image are filled by Poisson blending from correctly exposed LDR images using the homography defined around.

Sen et al. [14] proposes an approach aiming at constructing an HDR image which minimizes a two-term energy function after selecting the LDR image with the smallest number of under and overexposed pixels as the reference image. The first term of the energy function depends on the pixelwise difference between the HDR image and the reference image. The second term depends on the Multi-source Bidirectional Similarity Metric between the HDR image and the remaining LDR images based on Simakov et al. [15]. The resulting HDR image has a pixelwise similarity to the reference image. In order to improve the computational efficiency, a multiscale approach is used in this study.

Similar to Sen et al., after selecting the LDR image with the best exposure as the reference image, Hu et al. [16] constructs one latent image for each one the input LDR images which looks like the reference image but exposed as the input LDR image. While producing the latent images, an intensity mapping function is estimated. Other images in the input image set contribute to the latent images using the Generalized PatchMatch [17] process. HDR output image is constructed by merging the latent images.

Granados et al. [18] uses Markov Random Fields (MRF) to construct an HDR image from the input LDR images. They aim at minimizing the noise, visual discontinuity and ghosting artifacts with their proposed energy function.

In addition to the subjective evaluation of the authors, there is only one study on the comparison of the previously mentioned methods [19]. In Hadziabdic et al. [19], the four methods proposed by Sen et al, Zimmer et al., Photoshop and Photomatix programs are subject to the subjective evaluation of 30 participants. The findings indicate that Sen et al. produces better results than the other compared algorithms. The main contributions of our study may be listed as follows:

- Use of a new input data set with different levels of difficulty and different image properties which is open to improvement,
- Subjective evaluation of the results using different criteria,
- Comparison of the results obtained by an objective metric and subjective evaluation results, assessment of whether the subjective metric is suitable as a measure of the deghosting quality,
- Comparison of the method which is found as the superior in the previous study and another method which carries similar properties to it.

Table I: Properties of image sets used in the experiment according to scene motion (S), camera motion (K) and the type of moving object (C). +: Present, -: Not Present, P: Person, O: Object

LDR Image Set	**S**	**K**	**C**
Corridor	-	+	-
Cars (Handheld)	+	+	O
Cars (Tripod)	+	-	O
Library1 (Handheld)	-	+	-
Library1 (Tripod)	-	-	-
Library2 (Handheld)	+	+	P
Library2 (Tripod)	+	-	P
Office	-	-	-
Pedestrians (Handheld)	+	+	P
Pedestrians (Tripod)	+	-	P

II. THE EXPERIMENTAL FRAMEWORK

The data set used in the experiment consists of 10 LDR image sets captured using Canon EOS 550D camera with Magic Lantern 2.3 firmware. Each LDR image set consists of 9 images with 1 exposure value (EV) difference between subsequent LDR images. The image with exposure value of zero is chosen as the reference image for each one of the tested algorithms. All of the input images are resized to 1024×683. One of the image sets contains a completely static scene with no motion at all. The motivation behind including such a static scene is to see whether the tested algorithms introduce any additional visual artifact in scenes with potentially no ghosting artifact. Another image set contains only translational camera motion. Remaining image sets are captured using both tripod and handheld camera. The images are captured in both JPEG and camera RAW formats and they are a subset of a larger data set which is planned to be opened to public access of other researchers to test their algorithms. The properties of each image set are given in the Table I.

The MATLAB source codes provided by Sen et al. [14] and Hu et al. [16] are used in the experiment. For the objective quality measurement, the MATLAB source code provided by Liu et al. [20] is used. The study of Liu et al. [20] aims at objective evaluation of deblurring algorithms. Since artifacts introduced by HDR deghosting algorithms visually resemble the artifacts introduced by deblurring methods, the quality metric proposed by Liu et al. [20] is found suitable for the assessment of the HDR deghosting algorithms.

The HDR images with no deghosting are used as the blurry input images required by the Liu et al. For the Sen et al., the images obtained from camera RAW files with linear camera response curves are used as the inputs. In order to covert the camera RAW files to the desired format, open-sourced software called dcraw [21] is used. For visualization purposes, the Bracket [22] software is used. Since Hu et al. stated that their method does not require input images to have a linear camera response curve, the sRGB images obtained from the camera RAW files are used as inputs. In some of the overexposed and underexposed images with linear camera response curve, it is observed that Hu et al. was not able to produce a successful output since all image pixels were marked as outliers.

III. EXPERIMENTAL ANALYSIS

The results are compared under *visual*, *objective* and *subjective* perspectives.

Table II: Algorithm performance assessment using the quality metric proposed by Liu et al. [20] (a value closer to zero is better).

LDR Image Set	Hu	Sen
Corridor	-9,82	-12,01
Cars (Handheld)	-7,73	-7,88
Cars (Tripod)	-8,06	-8,34
Library1 (Handheld)	-10,11	-7,98
Library1 (Tripod)	-7,45	-7,84
Library2 (Handheld)	-8,68	-9,17
Library2 (Tripod)	-9,01	-9,21
Office	-6,45	-7,32
Pedestrians (Handheld)	-10,14	-10,06
Pedestrians (Tripod)	-10,06	-9,76
Average	-8,75	-8,96

A. Visual Analysis

Due to the limited space, results of the algorithms for only 2 scenes are provided in the Figure 2. Some of the regions are magnified to display the artifacts better. According to these results, it is observed that for some images, the information in high-frequency texture regions are lost due to oversmoothing by Hu et al. On the other hand, a large amount of noise is observed in the outputs of Sen et al. when the input images are taken with a high-ISO setting under inadequate lighting conditions, due to the absence of a spatial smoothing term in the energy function used. Full resolution versions of all results obtained from the data set is provided at http://www.ceng.metu.edu.tr/~tarhan/siu2014/siu2014.zip.

B. Objective Analysis

The performances of the compared algorithms according to the objective quality metric proposed by Liu et al. [20] is provided in the Table II. According to this metric, it is observed that the method proposed by Hu et al. is more successful.

C. Subjective Analysis

In order to make a comparison with the objective quality metric, the algorithm outputs are subjectively evaluated by the authors according to the noise, image distortion and deghosting performance. The evaluation results are provided in the Table III.

In the table, x/y expression implies that Sen et al.'s method is preferred x times and Hu et al.'s method is preferred y times in the pairwise comparison. According to the subjective evaluation results, Hu et al. deals with the noise better than Sen et al. while Sen et al. is able to preserve the object and texture integrity better than Hu et al. The deghosting quality of both algorithms are found relatively close to each other; however, since Hu et al. could lose some of the image details, Sen et al. was preferred over Hu et al. with respect to the deghosting quality.

IV. DISCUSSION AND CONCLUSION

The results obtained in this study shows that the proposed HDR construction algorithms have different performances. According to the visual analysis, it was observed that the algorithms were able to successfully eliminate the ghosting artifacts; however, while doing this they introduce some noise and texture smoothing artifacts. These findings are supported by the subjective analysis study. It is found that the relation between the subjective and objective evaluation results depends on the noise and loss of details in the output images. This finding may be related to the relatively large weight assigned to the noise features used in Liu et al.

It was observed that the presence of noise and the degree of smoothing has a significant effect on the evaluation results. In the future, introduction of a quality metric matching the subjective evaluation results with different feature weights based on the approach on Liu et al. on a larger data set and proposal of a better deghosting algorithm is planned.

Table III: Participant preferences in the user study according to Noise (G), Image Distortion (B), Deghosting Quality (K) criteria (one of the participants has given one points to each one of the compared results in case of a tie).

LDR Image Set	G	B	K
Corridor	0/4	3/1	4/0
Cars (Handheld)	0/4	4/1	3/2
Cars (Tripod)	0/4	4/1	4/1
Library1 (Handheld)	1/3	4/0	4/0
Library1 (Tripod)	0/4	4/1	2/3
Library2 (Handheld)	0/4	4/0	3/2
Library2 (Tripod)	0/4	3/1	4/0
Office	0/4	4/0	4/1
Pedestrians (Handheld)	1/4	4/0	4/1
Pedestrians (Tripod)	1/4	4/1	4/1
Total	3/39	38/6	36/11

REFERENCES

[1] Erik Reinhard, Greg Ward, Sumanta Pattanaik, and Paul Debevec, *High Dynamic Range Imaging: Acquisition, Display and Image-Based Lighting*, Morgan Kaufmann, San Francisco, second edition edition, 2010.

[2] Shree K Nayar and Tomoo Mitsunaga, "High dynamic range imaging: Spatially varying pixel exposures," in *Computer Vision and Pattern Recognition, 2000. Proceedings. IEEE Conference on*. IEEE, 2000, vol. 1, pp. 472–479.

[3] Jian Guo and Sameer Sonkusale, "A high dynamic range cmos image sensor for scientific imaging applications," *Sensors Journal, IEEE*, vol. 9, no. 10, pp. 1209–1218, 2009.

[4] Michael D Tocci, Chris Kiser, Nora Tocci, and Pradeep Sen, "A versatile hdr video production system," *ACM Transactions on Graphics (TOG)*, vol. 30, no. 4, pp. 41, 2011.

[5] Bernd Hoefflinger, Ed., *High-dynamic-range (HDR) vision: Microelectronics, image processing, computer graphics*, Springer Series in Advanced Microelectronics. Springer, Berlin, 2007.

[6] Peter J Burt and Raymond J Kolczynski, "Enhanced image capture through fusion," in *Computer Vision, 1993. Proceedings., Fourth International Conference on*. IEEE, 1993, pp. 173–182.

[7] S. Mann and R. W. Picard, "On being 'undigital' with digital cameras: Extending dynamic range by combining differently exposed pictures," Tech. Rep. 323, M.I.T. Media Lab Perceptual Computing Section, Boston, Massachusetts, 1994, also appears IS&T's 48th annual conference, Cambridge, Massachusetts, May 1995.

[8] Paul E. Debevec and Jitendra Malik, "Recovering high dynamic range radiance maps from photographs," in *Proceedings of the 24th Annual Conference on Computer Graphics and Interactive Techniques*, New York, NY, USA, 1997, SIGGRAPH '97, pp. 369–378, ACM Press/Addison-Wesley Publishing Co.

[9] Mark A Robertson, Sean Borman, and Robert L Stevenson, "Dynamic range improvement through multiple exposures," in *Image Processing, 1999. ICIP 99. Proceedings. 1999 International Conference on*. IEEE, 1999, vol. 3, pp. 159–163.

[10] Tomoo Mitsunaga and Shree K Nayar, "Radiometric self calibration," in *Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference on*. IEEE, 1999, vol. 1.

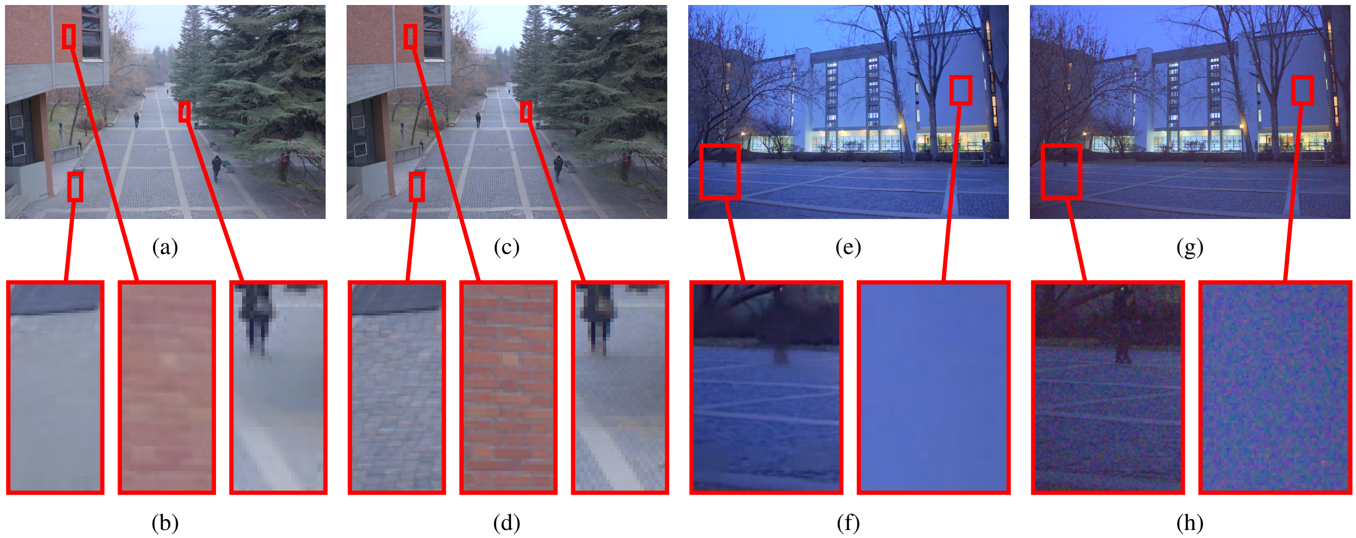

Figure 2: Outputs produced by the algorithms for two of the scenes. Some regions in the outputs of of Hu et al. in (a) and (e) are magnified in (b) and (f). Outputs of Sen et al. for the same images in (c) and (g) are magnified in (d) and (h).

[11] Henning Zimmer, Andrés Bruhn, and Joachim Weickert, "Freehand hdr imaging of moving scenes with simultaneous resolution enhancement," *Computer Graphics Forum*, vol. 30, no. 2, pp. 405–414, 2011.

[12] Jun Hu, Orazio Gallo, and Kari Pulli, "Exposure stacks of live scenes with hand-held cameras," in *Computer Vision–ECCV 2012*, pp. 499–512. Springer, 2012.

[13] Yoav HaCohen, Eli Shechtman, Dan B Goldman, and Dani Lischinski, "Non-rigid dense correspondence with applications for image enhancement," *ACM Transactions on Graphics*, vol. 30, no. 4, pp. 70, 2011.

[14] Pradeep Sen, Nima Khademi Kalantari, Maziar Yaesoubi, Soheil Darabi, Dan B Goldman, and Eli Shechtman, "Robust patch-based hdr reconstruction of dynamic scenes.," *ACM Trans. Graph.*, vol. 31, no. 6, pp. 203, 2012.

[15] Denis Simakov, Yaron Caspi, Eli Shechtman, and Michal Irani, "Summarizing visual data using bidirectional similarity," in *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*. IEEE, 2008, pp. 1–8.

[16] Jun Hu, Orazio Gallo, Kari Pulli, and Xiaobai Sun, "Hdr deghosting: How to deal with saturation?," in *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*. IEEE, 2013, pp. 1163–1170.

[17] Connelly Barnes, Eli Shechtman, Dan B Goldman, and Adam Finkelstein, "The generalized patchmatch correspondence algorithm," in *Computer Vision–ECCV 2010*, pp. 29–43. Springer, 2010.

[18] Miguel Granados, Kwang In Kim, James Tompkin, and Christian Theobalt, "Automatic noise modeling for ghost-free hdr reconstruction," *ACM Transactions on Graphics (TOG)*, vol. 32, no. 6, pp. 201, 2013.

[19] Kanita Hadziabdic, Jasminka Hasic Telalovic, and Rafal Mantiuk, "Comparison of deghosting algorithms for multi-exposure high dynamic range imaging," in *Spring Conference on Computer Graphics*, 2013.

[20] Yiming Liu, Jue Wang, Sunghyun Cho, Adam Finkelstein, and Szymon Rusinkiewicz, "A no-reference metric for evaluating the quality of motion deblurring," *ACM Transactions on Graphics (TOG)*, vol. 32, no. 6, pp. 175, 2013.

[21] D. Coffin, "Dcraw: Decoding raw digital photos in linux," http://cybercom.net/dcoffin/dcraw, last access: February 2014.

[22] A.O. Akyüz, "Bracket cross-platform high-dynamic-range (hdr) photo manager," http://www.ceng.metu.edu.tr/~akyuz/bracket/bracket.html, last access: February 2014.