Expert evaluation of deghosting algorithms for multi-exposure high dynamic range imaging

K. Karaduzovic-Hadziabdic[1] and J. Hasic Telalovic[1] and R. Mantiuk[2]

[1]International University of Sarajevo, Bosnia and Herzegovina
[2]Bangor University, United Kingdom

Abstract

In this work we analyze the results produced by high dynamic range (HDR) multi-exposure merging algorithms that handle motion artefacts. Within this analysis, we introduce the criteria for the evaluation of these algorithms. In order to perform a comprehensive evaluation we propose a dataset categorized into different types of scenes, each posing a challenge for the evaluated algorithm. For the analysis, we select and customize a tone-mapping operator (TMO), which minimizes a chance for TMO artefacts and allows us to inspect resulting HDR images on a conventional sRGB display. We show the results of five state-of-the-art algorithms evaluated based on the proposed criteria by performing expert evaluation analysis and draw a conclusion for each of these algorithms.

Categories and Subject Descriptors (according to ACM CCS): HDR imaging, deghosting algorithms

1. Introduction

Generating HDR images by merging a sequence of multiple low dynamic range (LDR) images [DM08] of dynamic scenes (often captured hand-held) results in motion artefacts. This poses a serious drawback of using the multi-exposure technique for the reconstruction of HDR in dynamic scenes. As a result, many algorithms have been developed to deal with removal of artefacts due to scene motion, camera motion or both. Srikantha et al. [SS12] provide a detail review of deghosting algorithms in HDR imaging. Since their review several more deghosting algorithms have been developed [SKY*12, HGPS13, GKTT13]. Therefore, it is desirable to analyze their performance and improvement over the older methods.

There are four main approaches reported in literature in handling motion artefacts: 1) algorithms that only use a subset of exposures for reconstruction of HDR of a scene [GGC*09]. These algorithms first detect ghosted pixels and then handle ghosting by excluding ghost detected regions from the composite HDR image. Such algorithms often have a reduced dynamic range of moving HDR content. 2) algorithms that perform alignment of different exposures before generating an HDR image [War03, TM07]. Algorithms based on the optical flow methods are very successful in aligning different exposures. Zimmer et al. [ZBW11] developed a state-of-the-art optical flow algorithm that aligns different exposures before merging to HDR image. 3) integration of alignment of LDR images and reconstruction of HDR image [SKY*12], 4) and recently, construction of ghost-free HDR images by performing noise modeling of the images proposed by Granados et al. [GKTT13].

Many algorithms have been developed to solve the deghosting problem of dynamic content in HDR imaging. All these algorithms show as result several scenes to demonstrate the success of their algorithm. None show their results on a wide range of scenes with varying complexity. In this paper we present a novel framework for evaluation of deghosting algorithms and introduce a criteria for their evaluation. In order to test the algorithms on a wide range of scenes with varying complexity, we propose a dataset of LDR images categorized into different types of scenes. Finally, we analyze five state-of-the-art algorithms based on the proposed framework: Photoshop, Photomatix, Sen et al. [SKY*12] (Sen2012), Jun Hu et al. [HGPS13] (HU2013) and Zimmer et al. [ZBW11] (ZIM2011). The abbreviation in brackets of latter three algorithms will be used in the rest of the paper to refer to each of the algorithms.

2. Evaluation of deghosting algorithms

Karaduzovic et al. [KHM13] assess deghosting algorithms by performing subjective psychophysical experiments that involved pairwise comparison of deghosting algorithms. The psychophysical analysis carried out included nine scenes and four algorithms. In this work, a more comprehensive dataset is used, and an additional Hu et al. [HGPS13], algorithm is evaluated. We observed that studies run with naive and untrained observers do not bring much insights into why HDR merging algorithms fail, what types of artefacts they produce and what could be the cause for them. Therefore, in this study we rely on an expert evaluation, in which three expert observers scrupulously inspect each image.

2.1. Criteria for the evaluation of deghosting algorithms

We present the criteria for deghosting algorithm assessment. A simple static scene captured by a hand held camera will often result in a ***blurry*** HDR image, which is caused by small misalignments of LDR images. Many deghosting algorithms have been developed to handle these artefacts. [War03,TM07]. Placing camera on a tripod could avoid most such blurring artefacts, though small misalignment could be caused by a photographer pressing controls on a camera, or even by the movement of a mechanical shutter. The alignment algorithms work very well for a small camera movement but they fail if the camera movement results in motion parallax.

Most real life scenes contain object(s) in motion, where movement is captured across different exposures. ***Motion artefacts*** are often visible as multiple appearances of the same object in different locations in the composite HDR image called 'ghosts'. These artefacts may also be manifested in the form of translucent 'ghosts' or may contain blotches throughout the moving object. Some algorithms handle both camera and scene motion [SKY*12, HGPS13].

There are algorithms that attempt to handle motion artefacts by considering only a subset of exposures in regions of scene where motion occurs. For such algorithms, if the object in motion contains HDR content then the dynamic range of the moving object is reduced, e.g. [GGC*09]. These algorithms cannot reproduce details in both bright and dark regions that contain large motion if HDR of a scene is large. Thus, next evaluation criterion is ***recovery of high dynamic range***, as HDR deghosting algorithms should aim to use all information in the exposures and reconstruct possibly large dynamic range of a scene.

Noise, which may appear in the reconstructed HDR image is another evaluation criterion. The magnitude of noise is mostly dependent on the selection of exposures for merging and on the weights used to average their pixel values. For example, if an algorithm assigns a higher weighting term to a shorter exposure, it will reduce some motion artefacts (due to shorter capture time), but it will increase noise is dark regions, which are affected by camera's read-out noise. Recently, Granados et. al [GKTT13] developed a deghosting algorithm that models image noise and produces the combined HDR image.

When only one or two color channels are clipped in all exposures but the remaining channel(s) contain(s) valid pixel data, the composite HDR image may contain ***color artefacts***. This can for example be observed as the sky turning pale white as the blue channel gets clipped and the resulting color is desaturated.

We also observed that some algorithms fail to reconstruct the image regions that were occluded by moving objects, resulting on ***merging artefacts***. Such regions are usually strongly blurred or replaced by a uniform color and usually surround moving objects. Ideally, a good deghosting algorithm should be broadly applicable and aim to produce an HDR image that greatly reduces, or completely eliminates the above mentioned artefacts.

2.2. Algorithms

We apply the proposed criteria by performing an expert evaluation analysis of five algorithms. Two of these algorithms are implemented in the commercially available software packages Photomatix Pro (version 4.2.6), and Photoshop CS5 Extended (version 12.0). Others are: [HGPS13], a patch-based algorithm that produces a registered stack from a sequence of misaligned images of dynamic scenes. The algorithm automatically selects an image with most well-exposed pixels to be the reference image. Sent et al. [SKY*12] is another state-of-the-art patch-based algorithm. The algorithm performs joint optimization of image alignment and HDR merging. (SEN2012) algorithm also requires the definition of a reference image, which needs to be defined by the user. Zimmer et al. [ZBW11] use state-of-the-art optical flow approach to register LDR exposures to the reference image before merging to HDR image. We did not use Granados et al. algorithm [GKTT13] in our analysis because the algorithm requires RAW images, which were not available in our dataset.

3. Dataset

When developing a deghosting algorithm, authors often focus on assessing the algorithm on particular set of scenes. Since algorithm performance is scene dependant, we present a structured dataset that involves challenging real life scenes for testing different deghosting algorithms. The Middlebury dataset designed for evaluation of optical flow algorithms, proposed by Baker et al. [BSL*11] was used as a guideline for creating our HDR dataset. This dataset is classified as: 1) complex motion, 2) fast and abrupt motion 3) high texture motion, 4) independently moving objects, 5) small motion displacement , 6) large motion displacement, (this also includes scenes where the object displacement is large such

that the object does not appear in all exposures), 7) large region of the scene changes 8) non rigid motion 9) occlusion, 10) multi-view scenes, and 11) night scenes.
We have made this data available for the research community at `http://projects.ius.edu.ba/ComputerGraphics/HDR`. The website also includes a document with detailed information on the camera settings of how the dynamic range of each scene was captured.

4. Tone-mapping

For the analysis, we customized a TMO based on the fast bilateral filter [DD02]. The TMO attempts to show the original high dynamic range images on a conventional display while minimizing a chance for introducing artefacts. The tone-mapping is not intended to produce the best looking images or reproduce an appearance. Instead, it reproduces details exactly as they were captured in the HDR images, and compresses low-frequency content to fit within a dynamic range of a display. To do so, the log-luminance of the original HDR images is decomposed into base and detail layers using an edge-preserving fast bilateral filter [DD02]. The detail layer is left unchanged while the base layer is scaled linearly (in the log domain) so that the 0.2th percentile is mapped to black and the 99.8th percentile is mapped to white. The log-luminance is converted back to the linear luminance and the color is reintroduced using the color correction formula from [RMH09] .

5. Experiments

We performed an expert evaluation analysis of five mentioned algorithms. The resolution of all images was rescaled to full high-definition (HD) resolution (1920x1080) so that images could be seen in the native resolution on a full-HD display. The sequences of LDR images for each scene in our dataset were processed by each algorithm. Photoshop, Photomatix and (SEN12) algorithms produce as a result merged HDR image. For (ZIM2011) and (Hu2013) algorithms we merged the computed registered LDR images into the HDR image using the Photomatix software. All five algorithms performed both image alignment and deghosting. Then, all the generated HDR images were processed by mentioned customized TMO, producing a .jpg image that can be viewed on an sRGB display. Expert evaluation analysis was then carried out as a rating experiment, where three expert observers viewed each tone mapped image and provided a rating for each criteria described earlier. Observers also noted down an overall comment for each scene.

6. Results and Discussion

The results of representative graphs obtained from the rating experiment are shown in Figure 1.

(SEN12):The final result is greatly impacted by the selection of the reference image.The algorithm cannot extend the dynamic range of large saturated regions of the reference by using information from other exposures (see conwy_beach in the dataset). In general, the algorithm produces images of high global contrast (dynamic range), with pure colors and low black level. The algorithm generates images with a higher amount of noise than other algorithms tested. It also struggles with non-rigid and high texture regions with motion if they containt HDR content. Usually, the algorithm successfully handles deghosting and occluded regions in the scene (see occlusion in the dataset and in Figure 2).

(HU2013): The algorithm selects a reference image and performs alignment of remaining exposures. We observed that the algorithm produces images of lower contrast (dynamic range) than (Sen2012). The black level is elevated. The algorithm sometimes distorts textures in the merged HDR image. (see tree_logs in the dataset and in Figure 3). Non-rigid scenes, rippling water surfaces with large portions of sun reflection contain visible artefacts (see iona_sun_beach_reflection in the dataset) and produce an unnatural HDR image. The algorithm is good in deghosting.

(ZIM11): Because Zimmer et al. [ZBW11] do not have their algorithm publicly available, we obtained registered exposures for representative scenes. The algorithm selects a reference image and computes optical flow for remaining exposures to register them with the reference. Being based on the optical flow, the success of LDR image registration greatly depends on accuracy of optical flow estimation. The algorithm works best with images with small object displacement (see green_plant_static_handheld in the dataset) and performs poorly for scenes with large object displacements.

Photomatix: We observed that the dynamic range of moving content is often reduced. This is possibly because a subset of exposures is used to handle moving objects. In other regions, information from all exposures is used to recover the dynamic range of a scene. The algorithm performs better than (Sen2012) in recovering high texture details in regions of high texture motion (see tree_logs in the dataset and in Figure 3). In general, Photomatix effectively reduces noise in merged HDR image. This is often at the expense of reduced dynamic range in these portions of the image. Photomatix performs clipping of highlights for some images (see conwy_bridge). This suggests that only a longer exposure was used for merging. We also observed that Photomatix produces images of lower contrast (dynamic range) than (SEN2012) and the black level is elevated. We found that Photoshop performs similarly for latter three observations. Regions affected by ghosts are handled well by the algorithm.

Photoshop: Similarly to Photomatix, the dynamic range of moving content is usually reduced and this could be because in this region the information from all the exposures is not used to recover the dynamic range of the scene. Photoshop performs better than Photomatix and Sen in recovery of details in high texture regions (see tree_logs in the dataset and Figure 3). The algorithm does not handle well regions affected by motion and produces visible ghosts in these regions.

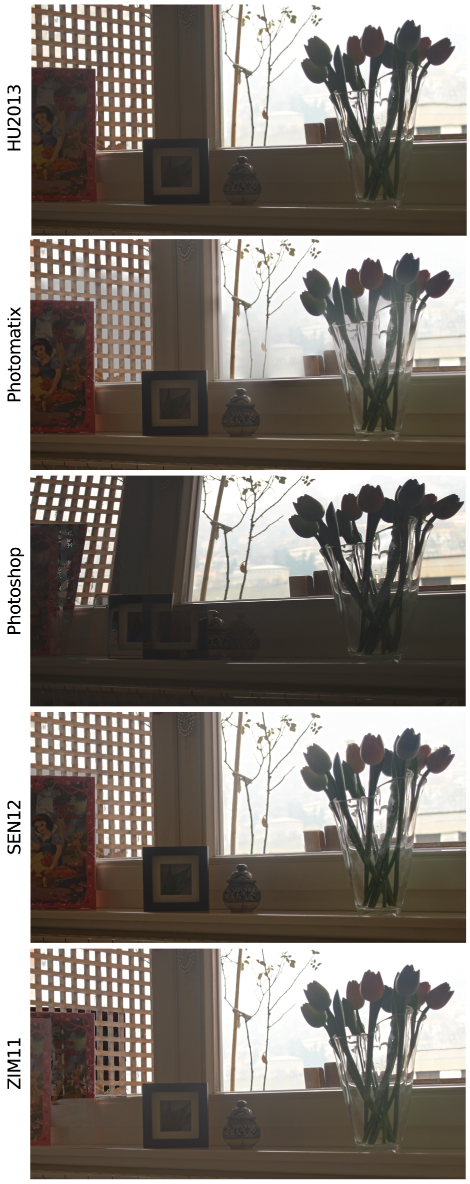

Figure 2: *The resuts of five tested algorithms on the 'occlusion' scene. Notice: ghosting in (ZIM11) and Photoshop; geometric distortions in Photoshop; blurry regions in the window in (HU2013) and Photomatix; distorted texture (building outside) in (HU2013)*

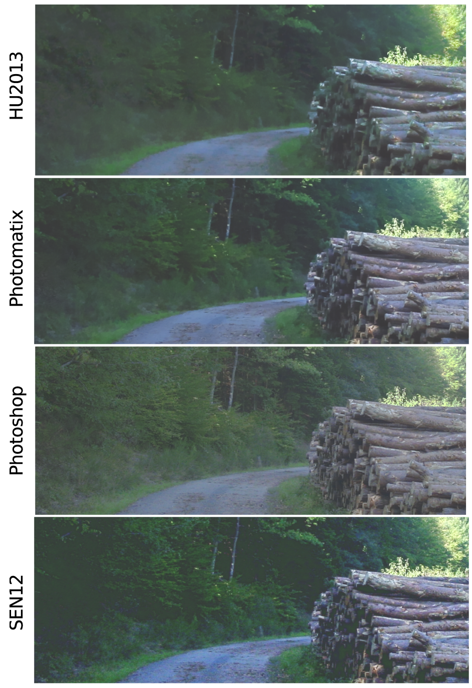

Figure 3: *The result of the four tested algorithms on the 'tree_logs' image (result for ZIM11 was unavailable). Notice: strong blurring in (HU2013) and Photomatix; contouring and noise in (SEN12).*

7. Conclusions

This paper presented a framework for analysis of deghosting algorithms in HDR imaging. Six criteria for evaluation of these algorithms are introduced: color artefacts, motion artefacts, blurring, noise artefacts, recovery of high dynamic range and merging artefacts. We also presented a comprehensive data set grouped into different scenes for evaluation of HDR deghosting algorithms. Five state-of-the-art algorithms were assessed by three experts and rating experiment was carried out to evaluate these algorithms based on the proposed criteria.

8. Acknowledgements

This work was partially supported by High Dynamic Range imaging COST Action IC1005. We would like to thank Jun Hu et al. [HGPS13] and Sen et al. [SKY*12] for making their code available to the public. We also thank Henning Zimmer for running his code on our datasets.

References

[BSL*11] Baker S., Scharstein D., Lewis J., Roth S., Black M. J., Szeliski R.: A database and evaluation methodology for optical flow. *International Journal of Computer Vision 92* (2011), 1–31. 2

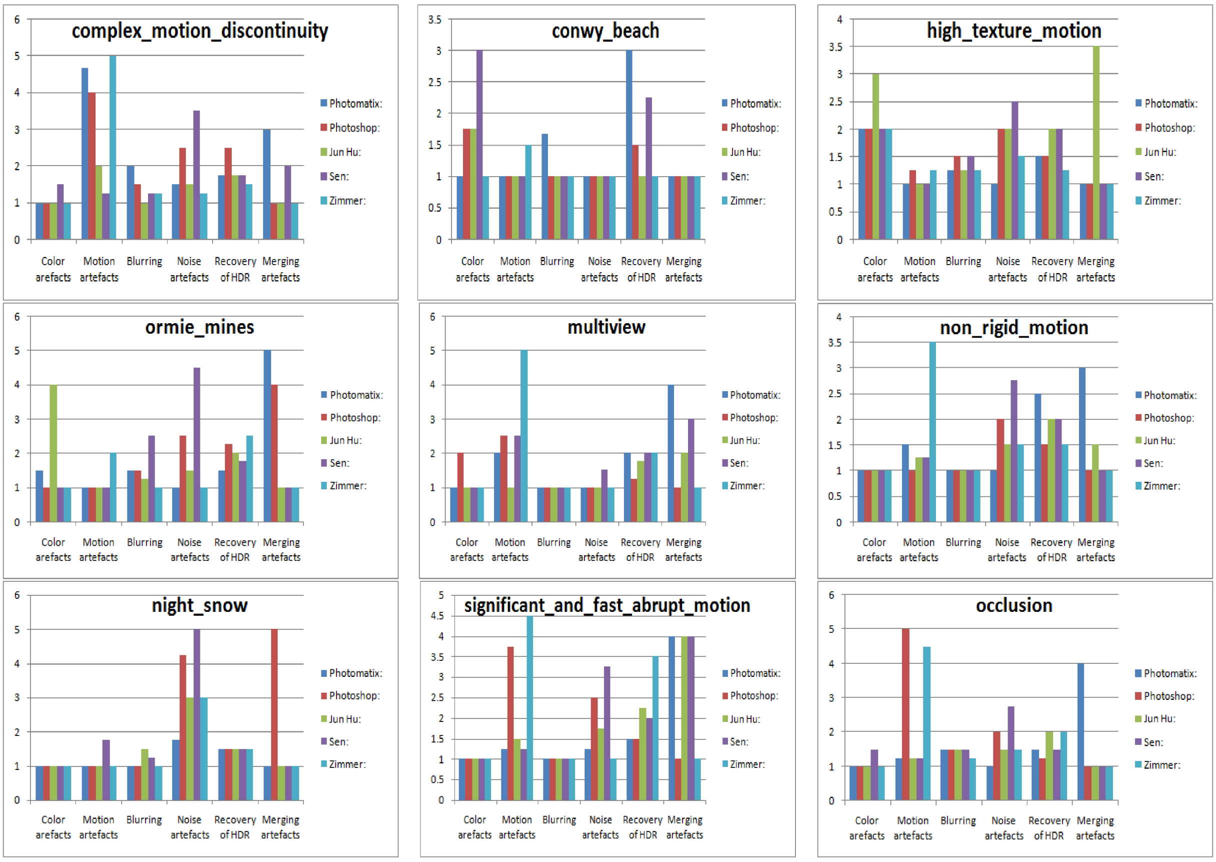

Figure 1: *Each graph displays performance of five deghosting algorithms for a single scene. The average rating values of three expert observers are shown on the y-axis: 1: invisible, 2: barely visible, 3: visible, 4: very visible, and 5: unacceptable.*

[DD02] DURAND F., DORSEY J.: Fast bilateral filtering for the display of high-dynamic-range images. *ACM Trans. Graph. 21*, 3 (July 2002), 257–266. 3

[DM08] DEBEVEC P. E., MALIK J.: Recovering high dynamic range radiance maps from photographs. In *ACM SIGGRAPH 2008 classes* (New York, NY, USA, 2008), SIGGRAPH '08, ACM, pp. 31:1–31:10. 1

[GGC*09] GALLO O., GELFAND N., CHEN W., TICO M., PULLI K.: Artifact-free high dynamic range imaging. *IEEE International Conference on Computational Photography (ICCP)* (April 2009). 1, 2

[GKTT13] GRANADOS M., KIM K. I., TOMPKIN J., THEOBALT C.: Automatic noise modeling for ghost-free hdr reconstruction. *ACM Trans. Graph. 32*, 6 (November 2013), 201:1–201:10. 1, 2

[HGPS13] HU J., GALLO O., PULLI K., SUN X.: Hdr deghosting: How to deal with saturation ? In *CVPR* (2013). 1, 2, 4

[KHM13] KARADUZOVIC K., HASIC J., MANTIUK R.: Comparison of deghosting algoirthms for multi-exposure high dynamic range imaging. In *In: Proc. of Spring Conference in Computer Graphics* (2013). 2

[RMH09] R. MANTIUK R. K. MANTIUK A. T., HEIDRICH W.: Color correction for tone mapping. In *In: Proc. of Eurographics, Computer Graphics Forum* (2009). 3

[SKY*12] SEN P., KALANTARI N. K., YAESOUBI M., DARABI S., GOLDMAN D. B., SHECHTMAN E.: Robust patch-based hdr reconstruction of dynamic scenes. *ACM Transactions on Graphics (TOG) (Proceedings of SIGGRAPH Asia 2012) 31*, 6 (2012), 203:1–203:11. 1, 2, 4

[SS12] SRIKANTHA A., SIDIBE D.: Ghost detection and removal for high dynamic range images: Recent advances. *Signal Processing: Image Communication 27*, 6 (2012), 650 – 662. 1

[TM07] TOMASZEWSKA A., MANTIUK R.: Image registration for multi-exposure high dynamic range image acquisition. In *In: Proc. of International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision* (2007). 1, 2

[War03] WARD G.: Fast, robust image registration for compositing high dynamic range photographs from handheld exposures. *JOURNAL OF GRAPHICS TOOLS 8* (2003), 17–30. 1, 2

[ZBW11] ZIMMER H., BRUHN A., WEICKERT J.: Freehand HDR imaging of moving scenes with simultaneous resolution enhancement. *Computer Graphics Forum (Proceedings of Eurographics) 30*, 2 (2011), 405–414. 1, 2, 3