# HDR IMAGE CONSTRUCTION FROM TRIFOCAL MULTIEXPOSURE IMAGES

*Alper Koz [a], Barış Demirkılıç [a,b], Yunus Bilge Kurt [a,c], Ahmet Oğuz Akyüz [a,d], Sinan Kalkan [a,d], A. Aydın Alatan [a,c], and Alan Chalmers [e]*

[a]Center for Image Analysis (OGAM), METU, Turkey, [b]Informatics Inst., METU, Turkey
[c]Dept. of Electrical and Electronics Eng., METU, Turkey, [d]Dept. of Computer Eng., METU, Turkey
[e]WMG, University of Warwick, United Kingdom

**ABSTRACT**

With the progress of autonomous vehicles, the sensing of the environment in more detail with higher dynamic ranges has become more important to classify surrounding objects and obstacles. While stereo HDR images for this purpose can provide advantages compared to the conventional LDR images, they suffer from limited dynamic ranges and spike-like noises due to the inaccuracies in disparity estimation. In this paper, we formulate the HDR image construction problem from trifocal multi-exposure images and develop a method which improves disparity estimation for better HDR image construction. Given the symmetric geometry of the trifocal setup, the proposed method uses the equivalence of disparities from middle to left and middle to right images to determine the reliable regions. The HDR radiance for the pixels in these reliable regions are estimated by using the weighted average of the warped images in different exposures and the middle image, whereas the radiance values outside the reliable regions are estimated by using only the middle image. The experiments with different exposure combinations for left, middle and right images reveal better performances of the proposed method compared to the stereo HDR imaging. It is also observed that the improvements are more apparent for larger disparities between the cameras.

***Index Terms***— High dynamic range, HDR, multi-exposure, disparity estimation, autonomous vehicles

## 1. INTRODUCTION

Autonomous vehicles (AVs) are increasingly becoming more pronounced owing to their potential to minimize human error. Leading companies such as Google, Volkswagen and Tesla take on this endeavor by investing significant amount of their resources. By the year of 2030, it is estimated that, with these efforts, 15% of the market will consist of autonomous cars [1]. AV's operational principle relies on sensing its environment by its sensors. Among the alternatives, ranging from radars to GPS receivers, optical cameras are very valuable as they can be used to perceive surrounding objects such as traffic signs, lights, and pedestrians to achieve fully autonomous or semi-autonomous driving.

Conventional cameras that are commonly used in AVs utilize low dynamic range (LDR) images of scenes. However, the real world scenes might contain high dynamic range (HDR) content which LDR images cannot sufficiently represent [2]. For reliable detection of the objects surrounding an AV in the visually and perceptually challenging conditions such as tunnel entrances & exits, sunsets & sunrises, and rainy & foggy weather conditions, HDR imaging has intrinsically more potential for safe driving.

The most common method in HDR imaging involves taking an exposure-varying image sequence of a static scene [3]. Since the response of a camera with respect to incident radiance is not linear, HDR methods recover the Camera Response Function (CRF) as the first step by using this image sequence. The images are then linearized using the CRF followed by the application of a tent-shaped weighting function to combine the radiance values of each exposure into a final HDR radiance map [3]. The final output is an HDR image which needs to be tone-mapped to be presented in conventional LDR display devices. Robertson et al. [4] use the similar steps but a few noise parameters are added in CRF recovery and HDR reconstruction. On the other hand, Mitsunaga and Nayar [5] assume a general CRF shape and compute the coefficients of a polynomial that models CRF. In contrast to these methods, exposure fusion algorithms [6] bypass CRF estimation, HDR reconstruction, and tone mapping steps to create a detail-rich LDR image by directly fusing the input exposures in the LDR domain.

Unfortunately, the real world scenes are dynamic, especially in driving scenarios. Therefore, using multi-exposed images captured from the same point of view usually produces ghosting artifacts in the generated HDR images [7]. An alternative that is especially practical in such settings is multi-view HDR imaging. In this case, multiple cameras with different viewpoints take images of a scene with different exposure settings at the same time. To merge these images, a reference view is chosen and the other non-reference view images are transformed onto the reference image by stereo-matching. One of the pioneering research efforts using this strategy is proposed by Lin and Chang [8]. The authors first calibrate the CRF by using Scale Invariant Feature Transform (SIFT) [9] between stereo images to find corresponding pixels. Then, stereo matching based on belief propagation developed by Sun et al. [10] is used to calculate disparity values. After this step, stereo-mismatches are identified by using a predetermined threshold for corresponding pixels' value difference. Similarly, Akhavan et al. [11] propose a framework for different methods to calculate disparity maps between stereo images with varied exposure settings and merge them into an HDR image. Batz et al. [12] extend the concepts for stereo images to the video. Apart from the stereo setups for HDR imaging, Bonnard et al. [13] use a special camera with 8 lenses to create a multi-view HDR image.

One of the biggest challenges on stereo HDR imaging is the stereo-matching step, which is prone to occlusions and exposure differences between different views. Another limitation is the use of only two views, which might be insufficient to capture the full dynamic range of a real scene. In contrast, trifocal HDR imaging might provide more dynamic range than the stereo-based HDR imaging due to the higher number of views with different exposures. Moreover, the geometric symmetry of such a setup for trifocal images can provide further advantages during disparity

**978-1-6654-3288-7/21/$31.00 ©2021 IEEE**

estimation as the disparities from the left and right images with respect to the middle image would be quite close to each other.

In this work, considering these advantages, we propose an HDR image construction method based on trifocal multi-exposure images. Despite the idea of multi-view HDR imaging is not novel, to our knowledge our study is the first one to perform an in-depth analysis of the advantages and challenges of a trifocal setup over a more traditional stereo setup. As the underlying idea of the proposed method, the symmetry between the middle-to-left and middle-to-right views is utilized to improve disparity estimation. More technically, a high confidence disparity map is obtained by using the geometric equality between the disparity maps from the left and right images. Then, the left and right images are warped to the middle view using this reliable disparity map followed by converting all images to the radiance domain by using the recovered CRF. For pixels with reliable disparities, the tent-shaped weighting function is used to calculate the weighted sum of the warped and the reference images. However, for pixels with unreliable disparities, only the reference middle image's radiance values are used. Finally, the HDR image is constructed as the combination of these two cases. The experiments reveal that the proposed trifocal HDR imaging gives better results in terms of HDR image quality metrics compared to the stereo HDR imaging. In addition, the quality improvement compared to the case without disparity checking is also verified.

In the next section, the HDR imaging problem from trifocal images is formulated. Section 3 gives the utilized datasets, which is followed by the description of the proposed method in Section 4. The experimental results are then presented in Section 5. Then, the paper is concluded in Section 6.

## 2. PROBLEM FORMULATION

Figure 1 illustrates a generic setup for a multi-view and multi-exposure camera system where $E_L$, $E_M$, and $E_R$ correspond to the exposures of the Left ($L$), Middle ($M$) and Right ($R$) cameras and $d$ is the distance between the optical centers of the cameras. In the case of unknown disparities in practice, the problem of trifocal HDR imaging first requires the estimation of the disparities from the middle image to the left image, and from the middle image to the right image. These operations can be expressed as,

$$D_{ML} = DE\left(I_M(E_M), I_L(E_L)\right) \quad (1)$$

$$D_{MR} = DE\left(I_M(E_M), I_R(E_R)\right) \quad (2)$$

where $D_{ML}$ and $D_{MR}$ refers to the estimated disparities from middle to right and from middle to left cameras, respectively. $DE(.,.)$ function corresponds to the disparity estimation operation. Note that the inputs are the right, left and middle images captured at the mentioned exposures.

Given the disparity maps, the left and right images are warped to the reference middle position. The HDR image is then constructed by using these warped images, the reference image at the middle position, and the estimated disparities,

$$H_{LRM} = f_{proposed}\left(I_L(E_L), I_M(E_M), I_R(E_R), D_{ML}, D_{MR}\right) \quad (3)$$

The problem of HDR image reconstruction from a trifocal camera involves the computation of the disparities as well as the design of the function, $f_{proposed}$, which takes the input exposures and the disparities as input. It then warps the input images, and merges them to produce an HDR image termed as $H_{LRM}$. The quality of the reconstructed image, $H_{LRM}$, should be first evaluated with respect

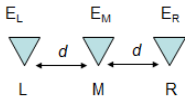
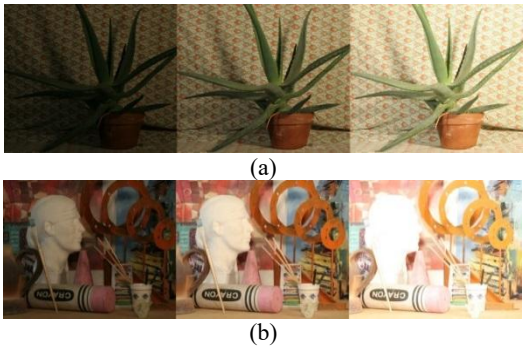



**Fig. 1.** A multi-view and multi-exposure camera setup



(a)

(b)

**Fig. 2.** The *Aloe* and *Art* multi-exposure and multi-view image sets from the Middlebury data set.

to the ground-truth HDR image rendered at the middle position by using the multi-exposure images taken at that position:

$$H_M = fusion\left(I_M(E_L), I_M(E_M), I_M(E_R)\right) \quad (4)$$

Also for the baseline comparison, the HDR image constructed with three images should be compared with HDR images constructed with only stereo pairs,

$$H_{ML} = f_{baseline}\left(I_M(E_M), I_L(E_L), D_{ML}\right) \quad (5)$$

$$H_{MR} = f_{baseline}\left(I_M(E_M), I_R(E_R), D_{MR}\right) \quad (6)$$

where $H_{ML}$ and $H_{MR}$ are the HDR images computed by using the middle-left and middle-right pairs, respectively.

## 3. THE DATASETS

*Aloe* and *Art* datasets of the Middlebury multi-view and multi-exposure image database are utilized for the experiments [14, 15]. Each dataset consists of 7 views (0..6) taken under three different illuminations (1..3) and with three different exposures (-1,0,1). The resolutions of the *Aloe* and *Art* datasets are 1282x1110 and 1390x1110 respectively.

All experiments are conducted using images with illumination 3. The middle-view image is fixed as View 3 with exposure (0) for all experiments. The left view is selected from (View 0, View 1, View 2) whereas the right view is selected from (View 4, View 5, View 6) based on the desired disparity between the cameras. Three different combinations of the exposure set [$E_L$ $E_M$ $E_R$], namely, [1 0 1], [-1 0 -1] and [-1 0 1], are tested in the experiments. Note that the left and right images have the same exposures in the first and second combinations, and different exposures in the third. Figure 2 illustrates the utilized images for the last combination as an example.

## 4. PROPOSED METHOD

The underlying idea of the proposed HDR image construction method from trifocal images is to use the symmetry of the trifocal setup. The absolute values of disparities of middle-to-right and middle-to-left in such a setup are proportional to the baseline distances. More specifically, if the baseline distance between the middle and left cameras is equal to the one between the middle and right cameras, then the disparity values for the same pixel in the

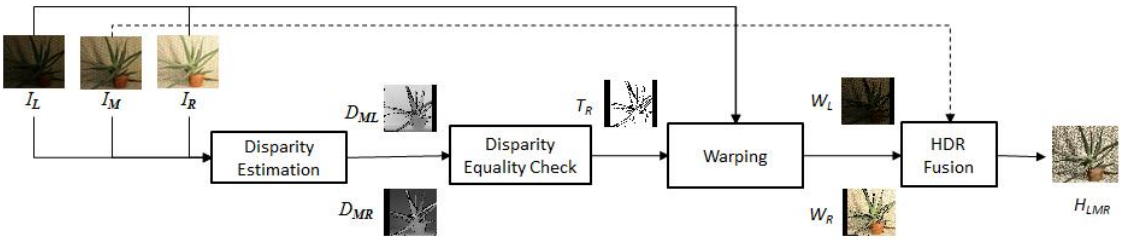



**Fig. 3**. The general scheme for the proposed HDR image construction from trifocal images

middle camera estimated for the left and right images would be equal to each other due the symmetry of the trifocal camera setup. Using this constraint, a reliable disparity map can be obtained to improve the constructed HDR image quality.

Given this fact, the first step of the proposed method is to determine the CRF for the reference view by using different exposure settings. This CRF could be used for both the disparity estimation and the HDR fusion steps. As the CRF is specific to a camera and does not change as long as the same camera settings are used, the recovery of the CRF can be done only once in a separate calibration phase during the installation of firmware of the vehicle.

After the CRF is obtained, the main stages of the proposed algorithm are as follows, as illustrated in Figure 3:

1. Disparity maps between the views are obtained by using any stereo-matching algorithm.
2. Disparity values are checked according to the symmetry. If the difference between the disparities are smaller than 1 pixel, then those pixels are assumed as reliable pixels for the warping operation.
3. Using the disparity maps, left and right views are warped to the middle view. In more detail, an image of zeros is first generated for the warped image to the middle position. The pixels of this image are filled by using the disparities from middle to left to warp the left image. The disparities are rounded to the nearest integer during warping. The same operations is also repeated for the right image.
4. The middle, the warped left and the warped right images are converted to the radiance domain using the already calibrated CRF.
5. For pixels corresponding to the reliable disparity values, a weighted sum of the radiance images is calculated by a tent-shaped weighting function [3]. However, in the untrusted disparity regions, only the reference middle image's radiance values are used. These untrusted regions also contain the occluded regions in the right and left images.
6. Finally, the HDR image is obtained as a combination of these values.

## 5. EXPERIMENTAL RESULTS

The experiments are performed for three different combinations of the exposure set [$E_L$ $E_M$ $E_R$], which are selected as [1 0 1], [-1 0 -1] and [-1 0 1], by using the *Middlebury* multi-view and multi-exposure data sets, Aloe and Art (Figure 2). Semi global matching (SGM) [16] algorithm is utilized for disparity estimation. HDR image metrics to compute the distance and similarity of the constructed and reference HDR images are selected as logarithm of the root mean square error (Log. RMSE) and the HDR visible difference predictor quality (HDR VDP Q) [17].

Figure 4 (a) and (b) gives the disparity maps estimated from middle to left and middle to right, namely, $D_{ML}$ and $D_{MR}$ as an example for the case of [-1 0 1]. Note that the contrast of the

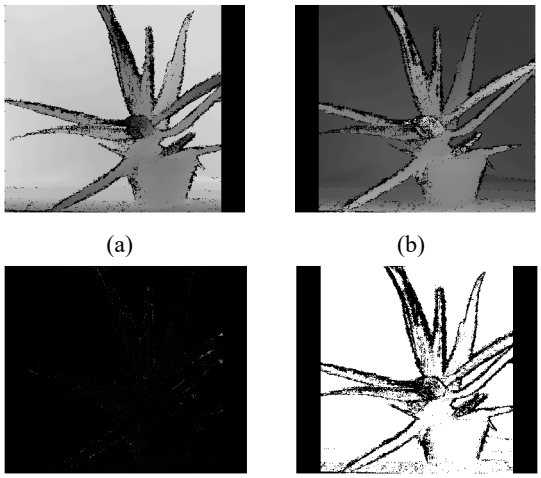



**Fig. 4**. The estimated disparity maps from (a) middle-to-right and (b) middle-to-left, (c) the absolute difference between the disparities, (d) the reliable regions for HDR radiance estimation by using warped images, where the disparities are close to each other

disparity images are opposite as the direction of the disparity estimations are opposite to each other. The absolute difference of the absolute disparities is illustrated in Figure 4 (c). While the overall difference is quite close to zero as indicated with dark regions, the difference at the edges are more apparent as those regions also correspond to different planes and occluded parts. This absolute difference is thresholded to indicate where the disparities are very close to each other as shown in Figure 4 (d). The white pixels indicate the trusted regions where both the middle image and the warped images are used to estimate the radiance values of the HDR image, whereas the radiance for the black regions are only estimated by using the middle image.

Table 1 gives the Log RMSE and HDR VDP Q results for the aforementioned exposure combinations. The results are given for stereo HDR images ($H_{LM}$ and $H_{MR}$) obtained by using (left and middle) and (right and middle) views and the proposed HDR image construction with trifocal images using disparity checking. In order to understand the improvement of the proposed method, the results are also presented without disparity checking. For all the three cases of exposures, the proposed method gives the minimum Log. RMSE as the distortions are decreased with the improved disparity estimation from both sides. The same observation is also valid for HDR VDP Q with an exception for the case of [-1 0 1] where $H_{LM}$ gives better HDR VDP quality. Given that the HDR VDP is more sensitive to spike-like noises, this suggests more analysis on the individual mapping of the pixels in warped images

**Table 1** Logarithmic Root Mean Square (Log RMSE) and HDR VDP Quality (HDR-VDP Q) results for the proposed trifocal HDR fusion methods with and without disparity check, and stereo HDR images.

| | | $H_{LM}$ (left and middle) | | $H_{MR}$ (right and middle) | | $H_{LRM}$ (left,middle,right) (without disparity check) | | $H_{LRM}$ (left,middle,right) (with disparity check) | |
|---|---|---|---|---|---|---|---|---|---|
| Image Set | Exposures [$E_L$ $E_M$ $E_R$] | Log RMSE | HDR VDP Q | Log RMSE | HDR VDP Q | Log RMSE | HDR VDP Q | Log RMSE | HDR VDP Q |
| Aloe | [ 1 0 1] | 0.139 | 9.40 | 0.160 | 9.23 | 0.109 | 9.58 | **0.108** | **9.64** |
| | [-1 0 -1] | 0.106 | 9.63 | 0.090 | 9.67 | 0.081 | 9.81 | **0.078** | **9.85** |
| | [-1 0 1] | 0.151 | **9.48** | 0.161 | 9.20 | 0.134 | 9.35 | **0.123** | 9.45 |
| Art | [ 1 0 1] | 0.150 | 9.42 | 0.160 | 9.421 | 0.108 | 9.68 | **0.092** | **9.80** |
| | [-1 0 -1] | 0.123 | 8.89 | 0.124 | 8.933 | 0.105 | 9.24 | **0.094** | **9.28** |
| | [-1 0 1] | 0.134 | 8.84 | 0.167 | 8.93 | 0.135 | 8.85 | **0.114** | **8.97** |

from both sides to form the HDR image.

Figure 5 illustrates the probability maps for the visual distortions in the generated HDR images for a sample case in terms of the HDR VDP metric. The color map increases from blue to green and green to red. The distortions, in particular on the edges of Aloe, are visible with higher values for the stereo HDR images $H_{LM}$ and $H_{MR}$. They decrease for the case of trifocal HDR images with lower values for the proposed method with disparity check.

As another aspect of the comparisons, Figure 6 gives the changes of Log RMSE and HDR VDP Q with respect to the disparity between the three cameras. Given the multiple-view set of *Middlebury* as (view0,..,view6), the cases with disparity 1, 2 and 3 correspond to (view2, view3, view4), (view1, view3, view5), and (view0, view3, view6) respectively. Compared to the previous experiments for the case of disparity 1, it is quite visible that the performance of the proposed method improves for higher disparities between the cameras. The selection of the disparity between cameras would be an important design parameter in further tests for both the performance of the camera and geometric design of the autonomous vehicles.

## 6. CONCLUSIONS

An HDR image construction method based on trifocal multi-exposure images is developed by using the equivalence of the disparities for the right and left images due to the geometric symmetry of the trifocal camera setup. The proposed method using trifocal images gives better performances than the stereo HDR images with more reliable disparity maps in terms of the HDR image metrics Log RMSE and HDR VDP Q. As the disparity between the cameras increases, it is also observed that the performance difference between trifocal HDR images and stereo HDR images increases as well. Future research will first focus on the collection of experimental data in different real-world driving scenarios with the proposed trinocular camera setup. Then, we will continue with the optimization of design parameters such as disparity and exposure combinations.


## ACKNOWLEDGMENT

This research is supported by the United Kingdom Royal Academy of Engineering under the Transforming Systems through Partnership Programme.

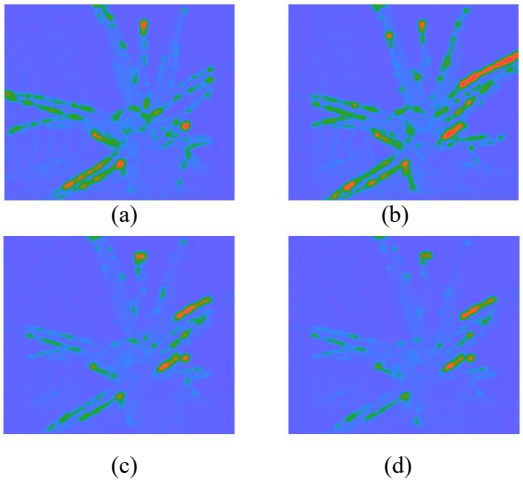



**Fig. 5**. HDR VDP distortion probality maps for (a) $H_{LM}$ , (b) $H_{MR}$, (c) $H_{LMR}$ without disparity check, (d) $H_{LMR}$ with disparity check.

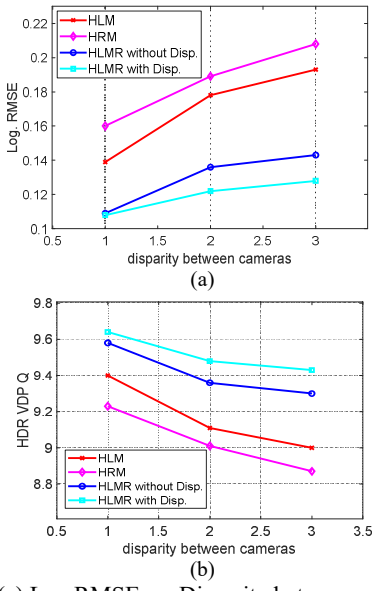



**Fig. 6.** (a) Log RMSE vs. Disparity between cameras and (b) HDR VDP Q vs. Disparity

## REFERENCES


[1] P. Gao, H. Kaas, D. Mohr, and D. Wee, "Disruptive trends that will transform the auto industry," McKinsey&Company, https://www.mckinsey.com/industries/automotive-and-assembly/our-insights/disruptive-trends-that-will-transform-the-auto-industry, January 2018.

[2] E. Reinhard, W. Heidrich, P. Debevec, S. Pattanaik, G. Ward, and K. Myszkowski, "High dynamic range imaging: acquisition, display, and image-based lighting," Morgan Kaufmann, 2010.

[3] P. E. Debevec and J. Malik, "Recovering high dynamic range radiance maps from photographs," in ACM SIGGRAPH 2008 classes , pp. 1-10, August 2008.

[4] M.A. Robertson, S. Borman, and R.L. Stevenson, "Estimation-theoretic approach to dynamic range enhancement using multiple exposures," Journal of electronic imaging, 12 (2), pp. 219-228, April 2003.

[5] T. Mitsunaga and S. K. Nayar, "Radiometric self calibration," in *Proceeding of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Vol. 1, pp. 374-380 Fort Collins, CO, USA, 1999.

[6] T. Mertens, J. Kautz, and F. Van Reeth, "Exposure fusion: A simple and practical alternative to high dynamic range photography," in Computer graphics forum , vol. 28, no. 1, pp. 161-171, Oxford, UK: Blackwell Publishing Ltd, March 2009.

[7] E. A. Khan, A. O. Akyuz, and E. Reinhard, "Ghost removal in high dynamic range images," in IEEE International Conference on Image Processing (pp. 2005-2008). October 2006.

[8] H.-Y. Lin and W.-Z. Chang, "High dynamic range imaging for stereoscopic scene representation," in 16th IEEE International Conference on Image Processing (ICIP), Cairo, Egypt, 2009, pp. 4305-4308.

[9] David G. Lowe, "Object recognition from local scale-invariant features," in Proceedings of the International Conference on Computer Vision. 2. pp. 1150–1157, 1999.

[10] J. Sun, N.-N. Zheng, and H.-Y. Shum, "Stereo matching using belief propagation," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 25, no. 7, pp. 787-800, July 2003.

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

[12] M. Bätz, T. Richter, J.-U. Garbas, A. Pabst, J. Seiler, A. Kaup, "High Dynamic Range Video Reconstruction from a Stereo Camera Setup", Signal Processing: Image Communi-cation, vol. 29, pp. 191-202, February 2014.

[13] J. Bonnard, C. Loscos, G. Valette, J. M. Nourrit, and L. Lucas, "High-dynamic range video acquisition with a multiview camera," in Optics, Photonics, and Digital Technologies for Multimedia Applications II , vol. 8436, p. 84360A, International Society for Optics and Photonics, April 2012.

[14] Middlebury Multiview Dataset, [Online]. Available: https://vision.middlebury.edu/stereo/data/.

[15] H. Hirschmüller and D. Scharstein, "Evaluation of cost functions for stereomatching," in IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2007), Minneapolis, MN, June 2007.

[16] H. Hirschmuller, "Stereo Processing by Semiglobal Matching and Mutual Information," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 30, no. 2, pp. 328-341, Feb. 2008.

[17] M. Narwaria, R. Mantiuk, M.P. Da Silva, and P. Le Callet, "HDR-VDP-2.2: a calibrated method for objective quality prediction of high-dynamic range and standard images," Journal of Electronic Imaging, 24(1), p.010501, 2015.