MDPI

Article

Use of an occlusion mask for veiling glare removal in HDR images

Federico Cozzi*, Carmine Elia, Giovanni Gerosa, Filippo Rocchetta, Matteo Lanaro and Alessandro Rizzi

Università degli Studi di Milano, Via Comelico 39/41, 20131, Milano (MI), matteo.lanaro@unimi.it
* Correspondence: federicocozzivm@gmail.com;

Version July 27, 2018 submitted to MDPI

Abstract: Optical systems in digital cameras present a limit during the acquisition of standard and High Dynamic Range Images (HDRI) due to the presence of veiling glare, an artifact caused by an unwanted spread of the source of light. In this paper we analyze the state of the art of veiling glare removal in HDRI, giving attention to the paper presented by Talvala et al.[5]. Then we describe an algorithm for veiling glare removal based on the same occlusion mask, to study the benefits provided by it in HDRi acquisition process. Finally, we demonstrate the efficiency of the occlusion mask method in veiling glare removal without any post production estimation and subtraction of it.

Keywords: HDR imaging; Veiling glare; Structured occlusion mask

1. Introduction

An HDRI (High Dynamic Range Image) is a digital image that represents a greater range of luminance levels than those obtained using a traditional sensor with a single exposure. HDRI acquisition techniques are based on a combination of several pictures of the same scene shot with different exposition parameters: doing that it's possible to include in a single image a larger amount of luminance values that in a Low Dynamic Range Image (LDRI) would be lost due to the quantization limit. This imaging technique is particularly used in science, photography[1] and medicine.

The capability of a camera to shoot an HDRI is limited by veiling glare[2]. Veiling glare is a global illumination effect caused by multiple light dispersions inside the camera's optical system. Veiling glare is different from the lens flare phenomenal: the second one appears when a punctiform and intense light source introduces in the image geometrical artifacts well defined; the veiling glare is instead related to extended sources of light and is similar to a blurred lens flare: it can be also caused by light sources out of the field of view but in this case it can be avoided using a lens hood.

Ideally a punctiform light source should illuminate a single pixel of the sensor, but in reality it also affects the pixels in the proximity, with an intensity proportional to the distance from the correct pixel and to the quality of the lens. The 2D function that describes the intensity of this phenomena is the Glare Spread Function (GSF) of the lens (Figure 1).

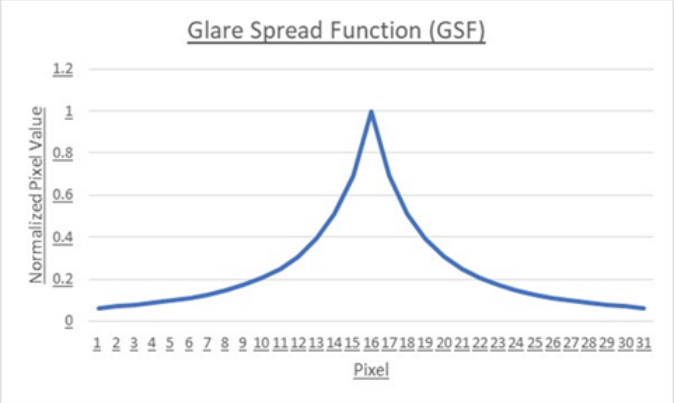


Figure 1. Example of Glare Spread Function (GSF).

In Section 2 will be presented the state of the art and some techniques of veiling glare removal during the acquisition and post-processing phases. In Section 3 will be described a technique of veiling glare removal based on spatial grid; in Section 4 will be presented the algorithm that we used to reconstruct the images, its implementation, testing and results. At last, in Section 5 will be discussed the limits and possible future developments of the technique proposed.

2. State of the art

Most of veiling glare removal techniques are focalized on refining optical elements of the system, for example improving the lenses coating [3]. Other methods are instead based on post-processing techniques, removing the veiling glare from images that are already affected by it. A common approach is using deconvolution techniques on the image, as presented in Starck et al.[4].

One deconvolution method for the veiling glare removal in post-processing is illustrated by Talvala et al.[5]: this technique is based on the hypothesis – unjustified- that the GSF is space-invariant and consecutively that exist only one GSF for the image; this GSF convoluted with the scene creates the image with the veiling glare. However, while trying to correct the image with the deconvolution we will incur in two main problems: first, the image parts characterized by the presence of glare become too much noisy due to the quantization limit; second, glare residual remains, characterized by a different chromatic component in various zones. Talvala et al.[5] conclusion is that in the image there isn't enough information for a successful veiling glare removal in post-production.

Nayar et al.[6] find as a correct approach for the veiling glare removal the separation between direct and global light, using a structured illumination and/or an occlusion mask (Figure 2). From this, Talvala et al.[5] develop their idea proposing the usage of an high frequency grid as occlusion mask. In the next section this method will be discussed and a variant of it will be proposed, in order to demonstrate the direct influence of the grid on the veiling glare.

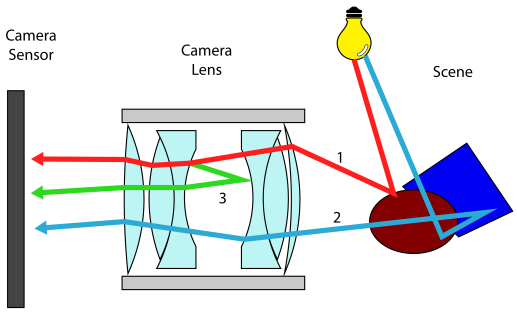


Figure 2. Subdivision of light in three different component[6]: direct light (blue) global light from the scene (red) and global light from the camera lens (green).

3. Occlusion Mask

Talvala et al.[5] placed a high-frequency occlusion mask between the camera and the scene to limit the presence of global light in the image (figure 3). Since glare is composed only by low spatial frequencies, they used a grid with an average occlusion factor of α to reduce the amount of veiling glare recorded in every pixel by a factor of α. It's necessary to shoot multiple captures, translating the mask in both directions (vertical and horizontal). This process is necessary to record only the direct component of light in all the scene. For a given α, at east $1/\alpha$ captures are needed to represent the entire scene through unoccluded regions. A small α involve in an elevated number of captures but also a lesser amount of glare and global illumination component. A large α grants a quick capture at the cost of increasing the noise. These are the reasons why a tradeoff is necessary. In their work, they create an HDR image for each grid position. Then, they estimate the amount of veiling glare in the occluded regions center. Next, a Gaussian filter is applied to spread the glare in the entire image. Finally, they subtract the estimated glare from the original image, and discard the regions outside the mask holes to realize a new glare-free image. The purpose of our project is to evaluate the effectiveness of the occlusion mask in an High Dynamic Range Image observing how veiling glare is reduced without an estimation and a subtraction, considering not occluded regions only.

4. Implementation

4.1. Acquisition

4.1.1. Grid and Materials

Based on the work of Talvala et al.[5] we used, as occlusion mask, an A3 matte black cardboard, laser cut as a grid with holes of 4x4mm and a 1cm period which leads to have an occlusion factor of α=0,16. This factor allow us to reconstruct the whole image. The grid has been positioned as near as possible to the scene (to make sure that, while focusing the scene, the grid wouldn't be too much blurred) with the help of a wooden structure as support. The grid has been horizontally moved 13 times, 1mm at a time, while for the vertical movements we used 4 thicknesses to raise the grid. Figure 3 shows the grid and the structures used in the experiment.

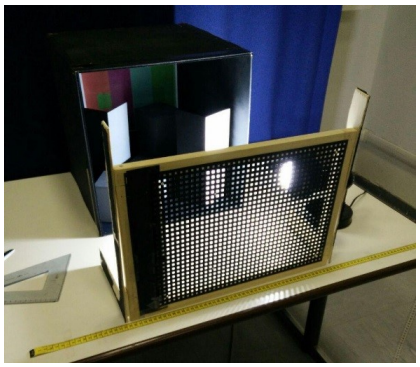


Figure 3. Structure and grid used in experiment.

4.1.2. Camera and Shot Settings

The final HDR image is composed by 7 images (with a range of exposition from -3 ev to +3 ev), each of them reconstructed starting from 13x4 images (13 horizontal steps x 4 vertical steps) of the scene occluded by the grid. In total we shot 364 photos. The pictures have been shot with a mirrorless Panasonic Lumix G7 with a m4/3 sensor of 16 Megapixels, and a 14-42mm lens (equivalent to a 28-84mm in a 35mm sensor) f/3.5-5.6. The characteristic curve of the optical system used is showed in Figure 4. The images are acquired with parameters shown in Table 1.

Table 1. Shooting Settings

	EV -3	EV -2	EV -1	EV 0	EV 1	EV 2	EV 3
FOCAL LENGTH	35mm (eq. 70mm)	35mm (eq. 70mm)	35mm (eq. 70mm)	35mm (eq. 70mm)	35mm (eq. 70mm)	35mm (eq. 70mm)	35mm (eq. 70mm)
ISO	400	400	400	400	400	400	400
F-STOP	8.0	8.0	8.0	8.0	8.0	8.0	8.0
SHUTTER SPEED	1/8	1/4	1/2	1	2	4	8

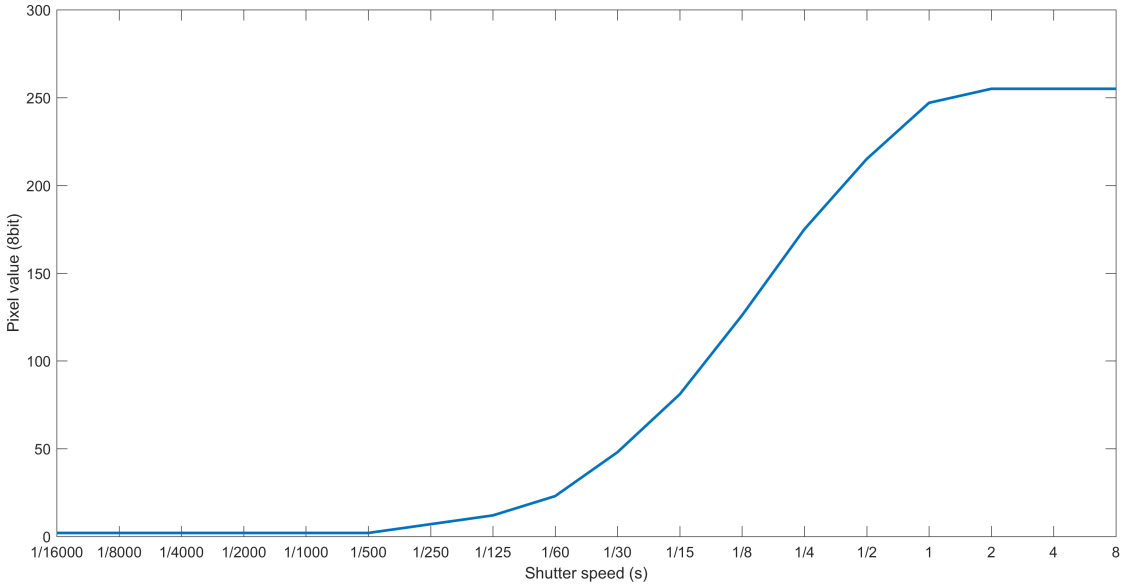


Figure 4. Characteristic curve of camera used (Lumix G7). To construct it we have photographed a white wall keeping an aperture of f8 and varying exposure times from about 1/16000 to 8s. We have then obtained the intensity for each photo, plotted obtained values on the Y-axis, and time on X.

	EV	cd/m^2
1	11.5	408.9
2	11.0	290.0
3	3.4	1.452
4	3.8	1.914
5	7.0	18.0
6	6.3	11.07
7	9.9	134.64
8	3.7	1.782
9	1.5	0.3948
10	6.5	12.69
11	5.1	4.815
12	3.4	1.452
13	2.8	0.9744
14	2.1	0.5992
15	1.4	0.3696
16	4.0	2.2
17	3.1	1.177
18	1.5	0.3948

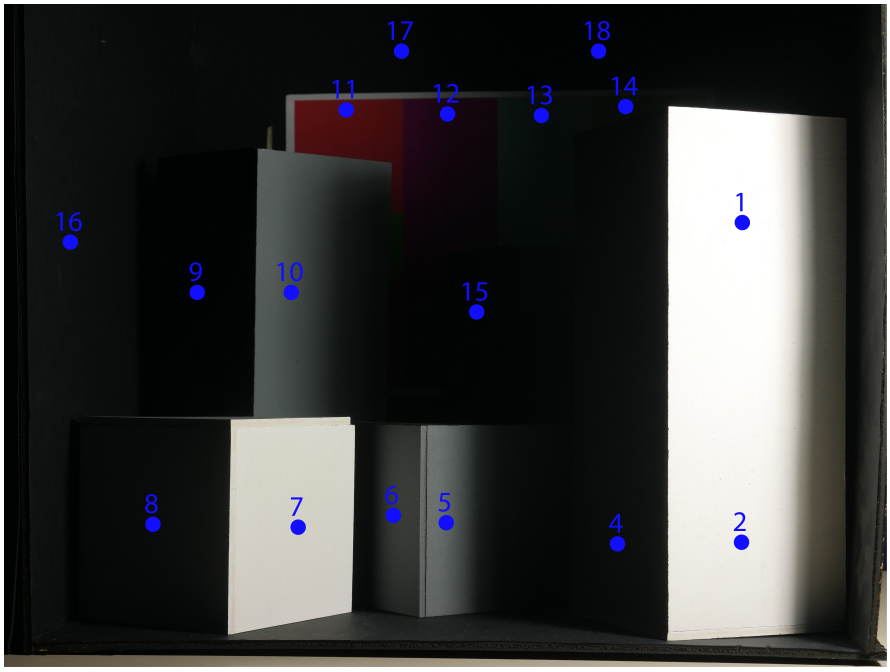


Figure 5. Luminance value record

The scene is composed by several cubes and parallelepipeds of various grey tones, inserted in a wooden box with black walls that can block the ambient light. The scene is illuminated with a lamp with a color temperature of 6400K expressly screened to not illuminate the grid, as it would cause more glare and a big amount of noise during the reconstruction phase (see after). On the background of the scene has been set an A4 sheet with different colored sections. Figure 5 shows the scene photographed with the different luminance values measured in several points.

4.2. Reconstruction Technique

At the end of the acquisition phase, the next step is to reconstruct the image of the scene assembling the areas not occluded by the mask. For each of the 7 exposures, we have taken a total of 52 photos (13x4): this means that we have collected, for each pixel P_{I} of the final image I 52, values. Reconstructing the image without an occlusive mask means assigning to each pixel P_{I} the value assumed when not occluded.

Let's call F_{P} the set of 52 values $f(P_{\mathrm{I}})$ assumed by every pixel P_{I}: we started considering the histogram of F_{P} to understand which values $f(P_{\mathrm{I}})$ represent the occluding mask and which ones represent the scene and must be used in the final reconstructed image. Ideally, every pixel should assume only two different values, one for the occluding mask and one for the scene: obviously, in a real context with the presence of non-sharp edges of the un-occluded zones, we expected to notice a bimodal distribution of the values of F_{P}, with the smaller mode as the occluded value, the bigger mode as the un-occluded one. On the base of that hypothesis, we firstly tried to exclude the occluded values choosing a fixed threshold, but it has not been possible to determine a single valid threshold for every pixel of the image. In the next attempts we tried to find a way to estimate a threshold adaptively according to the zone of the scene analyzed. To do that we made several reconstructions attempts:

1. Using as threshold the smaller mode and using as final value the mean of the remaining higher values (Figure 6a).
2. Using an arbitrary threshold and using as final value the mode of the remaining higher values (Figure 6b).
3. Using an arbitrary threshold and using as final value the mean of two or more mode of the remaining higher values (Figure 6c).

As can be seen in Figure 6, none of these attempts led us to an acceptable result. At last, we changed our approach: to be sure not to include the mask in the final image, we tried to consider only the higher values in F_{P}. Let's call:

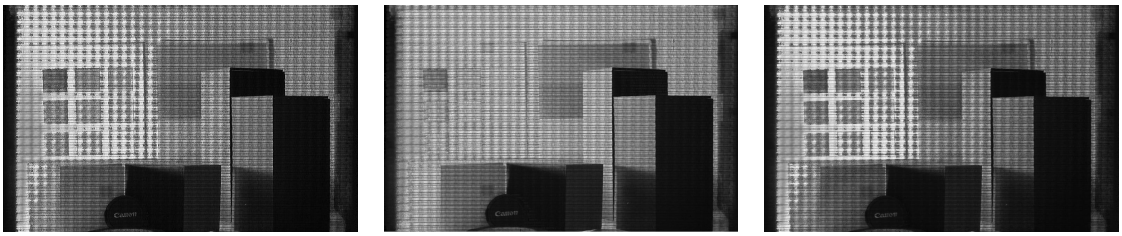
$$Q = \{unique\ ordered\ values\ in\ F_{\mathrm{P}}\}$$

$$m_1 = max(Q),\ m_2 = max(Q - \{m_1\}),\ m_3 = max(Q - \{m_1, m_2\});$$

$$Q_{\mathrm{M}} = \{m_1, m_2, m_3\}$$

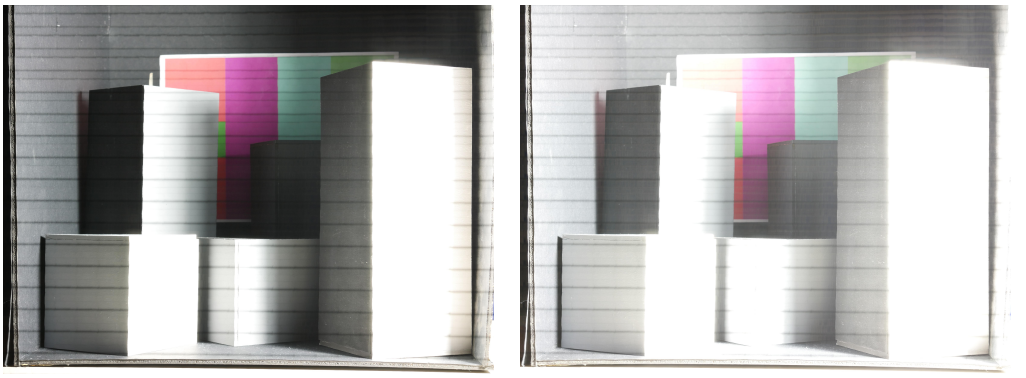
$$M = [f(P_{\mathrm{I}})\ in\ F_{\mathrm{P}} \mid f(P_{\mathrm{I}})\ \in\ Q_{\mathrm{M}}]$$

M is the set of the three highest different values assumed by the pixel (only if a pixel has more than two different values, that is cardinality $\#Q > 2$; otherwise M contains all the values in F_{P}. In the final reconstructed image, we use as value of every pixel P_{I}) the mean of its set of values A. The result produced by this method can be observed in Figure 7a, 7b, 7c: the remaining horizontal noise is caused by the poor precision of the vertical movement of the grid. Considering for the reconstruction only the higher values of each pixel leads to a slightly brighter image; it can also add more noise to the reconstructed image if the lamp is not properly masked and some direct light reaches the occlusion mask during the acquisition phase (the noisy image reconstructed in this case can be seen in Figure 7d): this condition has obviously to be avoided regardless from the reconstruction method used.



(a) **(b)** **(c)**

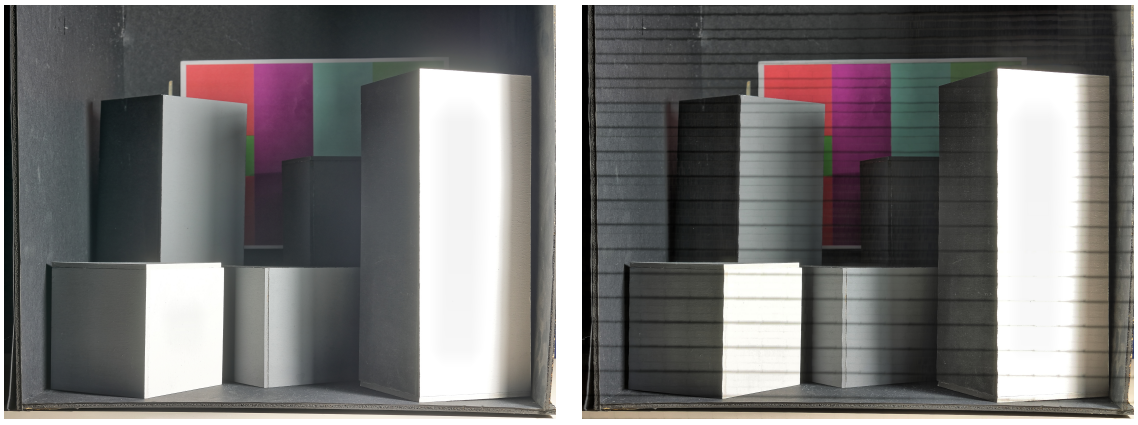
Figure 6. Results obtained from methods based on mode+mean (a) only mode (b), and mean of mode(c).



(a) **(b)**

(c) **(d)**

Figure 7. (a,b,c) : reconstruction of scene at EV 0, 2, -2; (d) : noise produced by direct light on the grid.



(a) **(b)**

Figure 8. Comparison between two HDR captures, tone mapped for display purpose. (a) : the veiling glare causes a loss of contrast, mostly in the darker regions; (b) : reconstructed image as result of the method presented in this paper, described in section 4.3.

4.3. The algorithm

Summing up, here are presented the principal steps of the algorithm to produce the final image:

1. Acquisition of an array of photos, to cover the entire scene with the holes of the occluding mask; this process is repeated for every exposure value of the final HDRI.
2. For every exposure, reconstruction of the scene starting from the un-occluded zones of the mask.
3. Assembling of the reconstructed images with different EV values, producing the final HDRI. We performed this operation using the software EasyHDR, which is based on the library DCRaw.

The final HDR image produced by this algorithm, confronted with the HDR image obtained without the occlusion mask method, is shown in Figure 8.

The difference between our procedure and the one used by Talvala et al.[5] lies in the order of the operations: Talvala et al.[5] mounted an HDRI for each mask translation (step 3) and then reconstructed the final image (step 2). They also used planar homology considerations in phase of reconstruction, while we reasoned on the histogram of values assumed by each pixel in pictures with the same exposure.

5. Results and Discussion

After the composition of the final HDR image we can see a visual improvement of the image quality: it appears clearer and less blurred comparing to both the HDRI with the original settings and an HDRI with the correct exposition.

The graph in Figure 9 shows the ratio between luminous intensity, measured in different points of the scene and the relative digit, normalized on the value of the most luminous point between the ones under consideration (not clipped). As we can see from the two curves, the one which represent the reconstructed HDR image is always under the other HDR image ‘s curve, which is the one with the original settings. From that we can say this method reduces the amount of veiling glare in the scene.

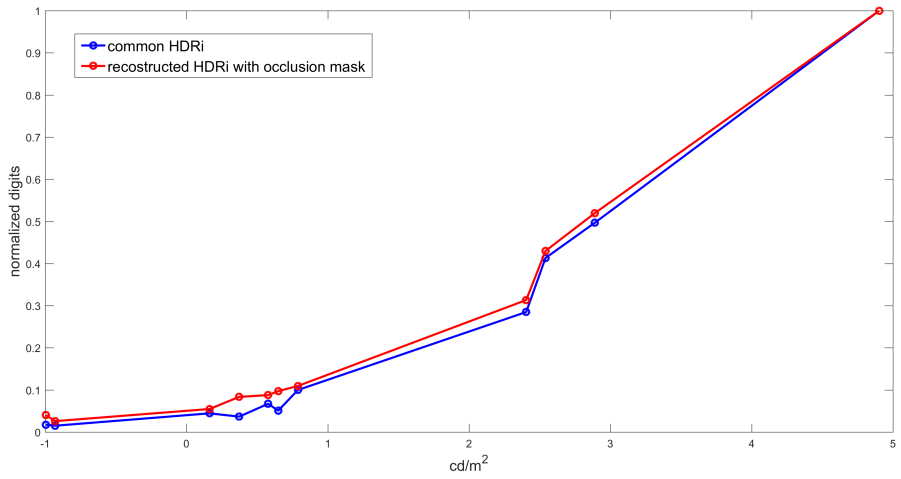


Figure 9. Ratio between luminous intensity and his corresponding digit for both curves

Furthermore, if we confront the two histograms, we can observe that the histogram of the reconstructed HDRI is more equalized. That confirms an improvement of the contrast and consequently of the the dynamic range.

Here we present a quantitative comparison between our method and the one proposed by Talvala et al.[5]. Glare and consequently its removal is image dependent, thus the presented comparison, done on different image sets, is mainly indicative. Talvala et al.[5] worked on a scene with a range of Luminance of 720:1 cd/m^2, while we measured in our scene a range of Luminance of 1106:1 cd/m^2. Regarding image dynamic range, Talvala et al.[5] could not estimate it in the whole scene after the process of reconstruction for the following reason: "At strong luminance edges where high-frequency glare is most apparent, we fail to recover scene content. Therefore, regions such as the top of the mirror or the blue egg become black after glare removal. Because of these artifacts, presenting whole-scene dynamic range figures would be inappropriate.". Then, they calculated the ratio of the values measured in some parts of the scene to the average value of the brightest zone (the background) before and after the reconstruction method. We calculated instead the Dynamic range of the two HDR images, with the occlusion mask (after the reconstruction process) and without it. The data we collected is shown in Table 2.

Table 2. Comparison of dynamic ranges measured on the HDRi with and without the occlusion Mask

	Without Occlusion Mask	**With Occlusion Mask**
Talvala et al. Object 1	560:1	22400:1
Talvala et al. Object 2	505:1	8850:1
Cozzi et al. Whole Scene	603:1	1503:1

6. Conclusion

In conclusion in this article we analysed the method of Talvala et al.[5] for the removing of the veiling glare in HDRI and we proposed the algorithm based on the same occlusion mask, to study the benefits provided by it in HDRi acquisition process. Based on the results obtained, we can say that using an occlusion mask during the acquisition of the image causes a global reduction of the amount of glare in the image, even without making any estimate or subtraction of it. Limitations of our study are due to the large number of shoots to reconstruct the whole scene and it's necessary a high precision, during the translation of grid, for getting low noise in reconstruction process. Negative aspect of our reconstruction alghoritm is given by the use of the highest pixel values for each shoot

taken with different shutter speed so reconstructed HDRI is averagely more luminous. In conclusion we can say that veiling glare can't be removed based only on data scene. We hope that in future will be proposed new and alternative methods.

References

1. REINHARD, E., WARD, G., PATTANAIK, S., AND DEBEVEC, P. High Dynamic Range Imaging - Acquisition, Display and Image-based Lighting. *Morgan Kaufman Publishers* **2006**, *500*, 142-149, DOI.
2. MCCANN, J. J., AND RIZZI, A. Veiling glare: the dynamic range limit of HDR images, In *Human Vision and Electronic Imaging*; SPIE, F., Editor2, A., Eds.; Publishing House: City, Country, 2007; pp. 32-58, ISBN.
3. PAUL A. BOYNTON, EDWARD F. KELLY Liquid-filled camera for the measurement of high-contrast images. iN*Proc.SPIE*; 5080; 2003; pp 50-100, ISBN.
4. STARCK, J., PANTIN, E., AND MURTAGH, F. Deconvolution in astronomy: A review. In *Publications of the Astronomical Society of the Pacific*; Editor1, F., Editor2, A., Eds.; Publishing House: City, Country, 2002; pp. 1051–1069, ISBN.
5. TALVALA, EV., ADAMS, A., HOROWITZ, M., LEVOY, M. Veiling Glare in High Dynamic Range Imaging. In *The Book Title*; ACM, F., Editor2, A., Eds.; Publishing House: City, Country, 2007; pp. 32-58, ISBN.
6. NAYAR, S. K., KRISHNAN, G., GROSSBERG, M. D., AND RASKAR, R. Fast separation of direct and global components of a scene using high frequency illumination. In *ACM Transactions on Graphics*; Editor1, F., Editor2, A., Eds.; ACM Press: New York, NY, USA, 2006; pp. 935–944, ISBN.