One-shot HDR Imaging via Stereo PFA Cameras

Tehreem Fatima[1[0000−0002−0146−1939]], Mara Pistellato[1[0000−0001−6273−290X]], Andrea Torsello[1[0000−0001−9189−4924]], and Filippo Bergamasco[1[0000−0001−6668−1556]]

[1]DAIS, Università Ca'Foscari Venezia. 155, via Torino, Venezia Italy
{tehreem.fatima, mara.pistellato, filippo.bergamasco, andrea.torsello}@unive.it

Abstract. High Dynamic Range (HDR) imaging techniques aim to increase the range of luminance values captured from a scene. The literature counts many approaches to get HDR images out of low-range camera sensors, however most of them rely on multiple acquisitions producing ghosting effects when moving objects are present.
In this paper we propose a novel HDR reconstruction method exploiting stereo Polarimetric Filter Array (PFA) cameras to simultaneously capture the scene with different polarized filters, producing intensity attenuations that can be related to the light polarization state. An additional linear polarizer is mounted in front of one of the two cameras, raising the degree of polarization of rays captured by the sensor. This leads to a larger attenuation range between channels regardless the scene lighting condition. By merging the data acquired by the two cameras, we can compute the actual light attenuation observed by a pixel at each channel and derive an equivalent exposure time, producing a HDR picture from a single polarimetric shot. The proposed technique results comparable to classic HDR approaches using multiple exposures, with the advantage of being a one-shot method.

Keywords: PFA Camera · HDR · Polarimetric Imaging

1 Introduction

The majority of modern digital cameras limit their capability to record the captured scene irradiance to 8 bit per channel. This translates in a significant weakness when acquiring high-contrast pictures, as some areas are going to be over or under saturated regardless the exposure time. Computational photography counts several techniques to recover High Dynamic Range (HDR) images, allowing for a better image visualisation and synthesis, as well as post-processing operations [18]. Classical approaches recover HDR scenes with standard camera sensors by merging multiple pictures taken with different exposure times [5, 19]. This limits their applicability to static scenes because moving subjects likely exhibit ghosting effects that must be properly accounted [23, 11]. Such limitation is not present in single-image approaches, but in this case obtaining an extended

dynamic range is more challenging [2, 3]. The recent advancement of learning-based methods allowed data-driven formulations of HDR reconstruction, usually exploiting the capabilities of Convolutional Neural Networks (CNNs). For example, Kalantari et al. [10] propose to reconstruct the HDR scene from a set of three exposures with moving subjects, while [15, 7] synthesise a set of images with different exposures starting from a single Low Dynamic Range (LDR) picture, recovering the HDR image in a standard way. Other alternative methods are based on encoder-decoder network architecture to recover the HDR image from saturated image regions [6, 30]. Finally, recent works as [25, 1, 16, 17, 12] propose alternative pipelines or architectures to extract missing information from LDR images and recover the original scene features.

In this work we propose a novel method to recover HDR images with a single shot with two Polarimetric Filter Array (PFA) cameras. A PFA camera mounts micro-polarizers at predefined orientations (0°, 45°, 90° and 135°) per macro-pixel, and thus allow the recovering of the polarization properties of incoming light. Since the polarization state is related to the features of the acquired scene, several applications are specifically designed to work with PFA sensors [22, 32, 31]. When placed in front of a camera, a linear polarizer attenuates the incoming light intensity according to the ray polarization state. Such effect is exploited to recover both angle and degree of polarization for each macro-pixel (see §2), but visually the intensity observed at each channel appears like an exposure time reduction. Our method is built upon this observation and is inspired by the work presented by Wu et al. [29]. Indeed, it is the first one proposing to use polarization properties of light to increase the dynamic range of their imaging system. Their major contribution is connecting light polarization state with virtual (or equivalent) exposure times, with an additional advantage of performing HDR reconstruction in a pixel-dependent way. Unfortunately, the resulting dynamic range depends on the Degree of Linear Polarization (DoLP) of the scene. Since the majority of real-world scenes are scarcely polarized, their technique is not as effective as taking multiple shots with an arbitrarily broad range of exposure times. To overcome such drawback, we propose a model that employs two PFA cameras capturing the scene at the same time as in a typical stereo steup, one of the two mounting an additional polarizer in front of the lens. Such additional filter allows the camera to receive exclusively highly polarized light, translating in a high dynamic range for the different channels, independently from the actual scene polarization. The actual angle and degree of polarization of the scene are recorded by the secondary camera (with no additional filter) and mapped into the image plane of the main camera. We propose a model to connect the values observed by the two cameras, so that we can recover the equivalent exposure times coherently with the original scene and the intensities observed by the main camera. Moreover, we propose a robust Stokes parameter estimation aimed at recovering missing values for over or under-saturated channels exploiting the redundancy of the camera filter angles. Experimental results show that our method offers a better HDR reconstruction with respect to the single camera setup, providing an image quality comparable with state-of-the-art techniques.

2 Polarimetric HDR

Together with amplitude and frequency, polarization is a property that is common to all type of vector waves. Being electromagnetic in nature, light makes no exception on this. Interestingly, such vector field consists of only two transverse components E_x, E_y, perpendicular to each other, that could be chosen for convenience to propagate in the z direction. This allows the complete description of E in terms of 4 scalar quantities called the Stokes polarization parameters $\begin{pmatrix} S_0 & S_1 & S_2 & S_3 \end{pmatrix}$ [9]. In this model, S_0 is the total intensity of the light, as we commonly capture with digital grayscale cameras. S_1 and S_2 express the *linear polarization* with respect to two reference frames mutually tilted by 45°. Finally, S_3 quantifies the amount of left-right *circular polarization* of the beam.
As a consequence of the Schwarz's inequality, Stokes parameter are related by the following formula:

$$S_0^2 \geq S_1^2 + S_2^2 + S_3^2 \tag{1}$$

showing that the total amount of light radiation S_0 is generally composed by a mixture of *polarized* (either linear and circular) and *unpolarized* light. In particular, when light is fully polarized Eq.1 becomes an equality, and when is fully unpolarized, the right-hand side is zero [9].

Stokes parameters are ubiquitous in their usage because they are directly measurable using two optical elements: *retarders* and *polarizers*. The former produces a phase-shift between E_x and E_y whereas the latter filters everything except the light radiation aligned with a certain transmission axis. This is exploited in PFA cameras to extract the first 3 Stokes parameters similar to how colour cameras use Bayer filter to recover light frequency. Pixels are masked with a repeated pattern of *linear polarizers* oriented at $0°, 45°, 90°$, and $135°$ wrt. the horizontal direction of the image. When properly demosaiced [20], we obtain four images $I_0, I_{45}, I_{90}, I_{135}$ grouping all the pixels with a certain polarizer orientation. From that, the first three Stokes parameters are simply computed as follows:

$$S_0 = I_0 + I_{90} = I_{45} + I_{135} \tag{2}$$
$$S_1 = I_0 - I_{90} \tag{3}$$
$$S_2 = I_{45} - I_{135}. \tag{4}$$

Unfortunately, S_3 cannot be obtained without a retarder, but circular polarization is relatively rare in nature [4]. Therefore, many applications simply assume that light is only composed by either unpolarized or linearly polarized light with a certain angle. To physically describe this behaviour, the Degree of Linear Polarization (DoLP) and the Angle of Linear Polarization (AoLP) are computed from Stokes parameters as:

$$\text{DoLP} = \frac{\sqrt{S_1^2 + S_2^2}}{S_0}, \quad \text{AoLP} = \frac{1}{2} \arctan \frac{S_2}{S_1}. \tag{5}$$

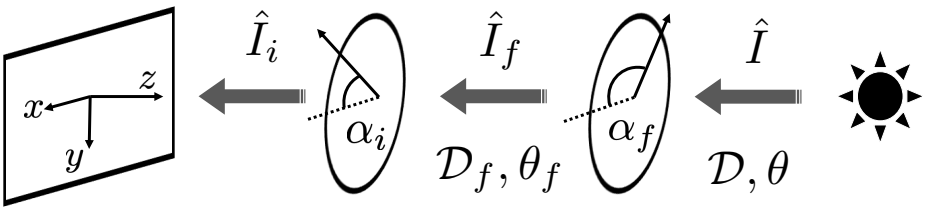

Fig. 1. Our proposed HDR imaging device: two linear polarizers are placed in the optical path from scene to sensor. The innermost is part of the PFA camera and oriented at a fixed angle $\alpha_i \in \{0, \frac{\pi}{4}, \frac{\pi}{2}, \frac{3\pi}{4}\}$, following the array pattern arrangement. The other one is placed outside the lenses and can be freely oriented with an angle α_f with respect to the camera horizontal axis x. The relative rotation of the two allows us to adjust the dynamic range of the system regardless the scene DoLP $\mathcal{D}$ and AoLP θ.

2.1 The proposed acquisition model

Let the image *irradiance* $\hat{I}$ be the amount of energy flowing from the scene to the sensor. When an (ideal) linear polarizer is placed in front of the sensor, as in a PFA camera, the irradiance is attenuated as follows:

$$\hat{I}' = \frac{1}{2}\hat{I}\big(1 + \mathcal{D}\cos(2\theta - 2\alpha)\big) \tag{6}$$

where $\mathcal{D}$ and θ are the DoLP and AoLP of the light entering the filter, and α is the orientation of the filter with respect to the camera horizontal axis. As we will see later, this attenuation has the same effect of reducing the the camera shutter speed (darkening the image) so it can be used as in classical HDR imaging.

From Eq.6 it is easy to note that the attenuation varies with the relative difference between the incoming AoLP and filter orientation α. $\hat{I}'$ is maximum when $\theta \cong \alpha(\mod \pi)$ and minimum when $\theta \cong \alpha + \frac{\pi}{2}(\mod \pi)$. The ratio between $\hat{I}'_{\max}$ and $\hat{I}'_{\min}$ (i.e. the dynamic range of the system) depends on the amount of linear polarization $\mathcal{D}$ of the light. Unfortunately, in a typical scene without a massive amount of specular reflections, the DoLP is low and inversely proportional to the surface's albedo due to the so-called *Umov's* effect [27, 14]. So, regardless the filters orientation, an HDR system using a built-in PFA is not particularly effective, because bright objects (the ones requiring a greater extent in filter attenuation) are characterized by a DoLP $\ll 1$.

Our idea is to use an additional linear polarizer, as sketched in Figure 1. In this way, light rays with intensity $\hat{I}$, DoLP $\mathcal{D}$, and AoLP θ are attenuated by two filters before being measured by the sensor. The first filter is placed in front of the camera lens and can be freely oriented with an angle α_f. The resulting intensity $\hat{I}_f$ is modelled by Eq.6 as:

$$\hat{I}_f = \frac{1}{2}\hat{I}\big(1 + \mathcal{D}\cos(2\theta - 2\alpha_f)\big) \tag{7}$$

The DoLP and AoLP after the filter should ideally be $\mathcal{D}_f = 1$ and $\theta_f = \alpha_f$, respectively. However, since the front polarizer is not perfect, we assume a certain

variability on that and model them explicitly. Then, light rays reach a second filter embedded in the camera's PFA. Depending on the pixel, such filter can assume only 4 possible angles α_i ($0 \leq i \leq 3$), where $\alpha_0 = 0$, $\alpha_1 = \frac{\pi}{4}$, $\alpha_2 = \frac{\pi}{2}$, $\alpha_3 = \frac{3\pi}{4}$. By substituting Eq.7 into Eq.6 we can model the irradiance $\hat{I}_i$ after the combination of both the filters:

$$\hat{I}_i = \frac{1}{4}\hat{I}\big(1 + \mathcal{D}\cos(2\theta - 2\alpha_f)\big)\big(1 + \mathcal{D}_f \cos(2\theta_f - 2\alpha_i)\big). \tag{8}$$

For the so called *reciprocity property* [5], the actual intensity values captured by the camera I_i are related to irradiance $\hat{I}_i$ and exposure time t_c through $g(I_i) = \hat{I}_i t_c$, where $g()$ is the *Inverse Camera Response Function* (ICRF). Algebraically, we can substitute the product $\hat{I}_i t_c$ with $\hat{I} t_i$ where:

$$t_i = \frac{1}{4}t_c\big(1 + \mathcal{D}\cos(2\theta - 2\alpha_f)\big)\big(1 + \mathcal{D}_f \cos(2\theta_f - 2\alpha_i)\big). \tag{9}$$

In other words, the four images $\hat{I}_0, \ldots, \hat{I}_3$ obtained by rotating the internal polarizer are de-facto identical to the four images that would be obtained without the filter but by changing the exposure time according to Eq.9. We call $t_0 \ldots t_3$ *equivalent exposure times*.

2.2 HDR imaging procedure

The equivalent exposure times for each pixel depends by the interplay between α_f, α_i and the scene polarization state $\mathcal{D}, \theta$. However, $\mathcal{D}$ and θ cannot be observed with our two-filter configuration because the external filter masks the full scene polarization to the internal one used for measuring. Therefore, we propose to use a second PFA camera (without any additional filter) to map the scene DoLP and AoLP to the former image.

The whole procedure can be summarized as follows:

1. Two PFA cameras are mounted side-by-side like in a typical stereo configuration, with a minimal baseline and hardware trigger to ensure that images are captured at the same time. Let *Cam0* be the one with an additional linear polarizer attached in front of the lenses and *Cam1* the other.
2. A picture is taken with both cameras at the same time. Let t_c be the *Cam0* shutter speed. *Cam1* shutter speed is not important as long as the image is exposed as good as possible[1].
3. Both the images are demosaiced to obtain $I^0_0, I^0_{45}, I^0_{90}, I^0_{135}$ and $I^1_0, I^1_{45}, I^1_{90}, I^1_{135}$ (subscript denote CCD filter orientation, superscript the camera number, and images are assumed to be normalized in $0 \ldots 1$ range after applying the ICRF estimated a-priori). The first three Stokes parameters are computed using Eq. 2, 3, 4.
4. Optical flow f mapping pixels from *Cam1* to *Cam0* is computed from S^0_0 and S^1_0 as described in [8].

[1] Considering that we are acquiring a scene with a high dynamic range, it will be unavoidable to over- or under-expose some areas.

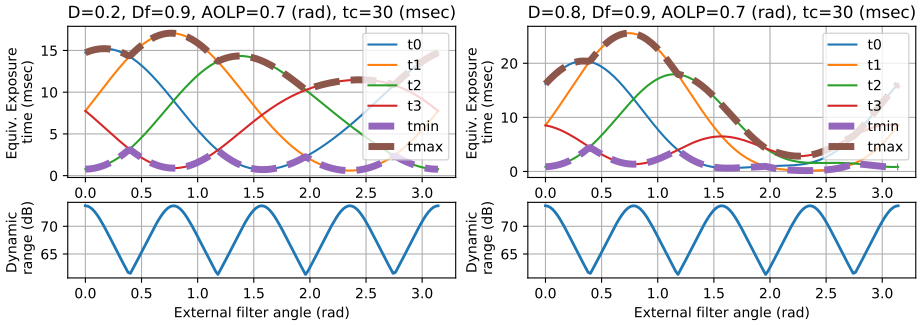

Fig. 2. Numerical examples of the equivalent exposure times (top) and dynamic range (bottom) for ideal scenes with a constant DoLP $D = 0.2$ (first column) and $D = 0.8$ (right column) varying the external filter angle α_f.

5. Scene DoLP and AoLP are computed from the Stokes parameters of *Cam1* (S_0^1, S_1^1, S_2^1) using Eq. 5. The resulting images are warped with f to obtain $\mathcal{D}$ and θ as it would have been observed from *Cam0*.
6. $\mathcal{D}_f$ and θ_f are computed from the *Cam0* Stokes parameters using Eq. 5.
7. The equivalent exposure times t_0, t_1, t_2, t_3 of $I_0^1, I_{45}^1, I_{90}^1, I_{135}^1$ are computed as in Eq. 9. Note that, unlike typical HDR techniques, exposure times are computed per-pixel as they depend on light polarization state that vary along the image.

The result is a set of 4 images $I_0^1, I_{45}^1, I_{90}^1, I_{135}^1$ as it would be obtained by bracketing the real shutter speed at times $t_0 \dots t_3$. Albeit any method can be chosen at this point to obtain the final HDR image, we use [5] for its simplicity and overall quality. It is essentially based on averaging each image with a Gaussian weight $w(l) = \exp\left(-\frac{(l-0.5)^2}{2\sigma^2}\right)$ giving more importance to irradiance values closer to the middle of the response function[2]. The final HDR image is thus:

$$I_{\text{HDR}} = \frac{w(I_0^1)\frac{I_0^1}{t_0} + w(I_{45}^1)\frac{I_{45}^1}{t_1} + w(I_{90}^1)\frac{I_{90}^1}{t_2} + w(I_{135}^1)\frac{I_{135}^1}{t_3}}{w(I_0^1) + w(I_{45}^1) + w(I_{90}^1) + w(I_{135}^1)}. \tag{10}$$

By manually rotating the external filter (i.e. varying the angle α_f) the dynamic range of the acquisition system can be adjusted to match the imaged scene. In Fig. 2 we show how the equivalent exposure times, computed as in Eq. 9, vary depending on α_f. If the scene is mostly unpolarized (first column), $t_{\min}$ and $t_{\max}$ will be loosely affected by the external filter angle. On the contrary, if the scene is polarized, the absolute range of equivalent times will be smaller when the filter is orthogonal to the scene AoLP. In both the cases, the dynamic range $\text{DR} = 20 \log \frac{255\, t_{\max}}{t_{\min}}$ vary from a minimum of ≈ 61 dB to a maximum of ≈ 74 dB (considering 8 bits per pixel images) regardless the actual scene polarization. In particular, when α_f is aligned with any α_i, we observe the maximum

[2] We empirically observed that $\sigma = 0.2$ usually gives satisfactory results.

DR and two out of the four images will have an equivalent exposure time equal to $0.5(t_{\max} + t_{\min})$. When α_f is in between two α_i (for example $\alpha_f = \frac{\pi}{8}$) we get the minimum dynamic range and all equivalent exposure times are equal to either $t_{\min}$ or $t_{\max}$. The important thing to notice is that such behaviour is independent by the actual scene polarization state. Indeed, the DR is only affected by: (i) the external filter angle α_f and (ii) the ability of the external filter to block all the unpolarized light (ie. the DoLP $\mathcal{D}_f$ of light exiting the filter). This flexibility is the main advantage of our proposed method.

2.3 Robust estimation of Stokes parameters

The accuracy of $\mathcal{D}$ and θ is crucial to compute the equivalent exposure times especially when α_i and α_f are almost aligned. Certainly, camera thermal noise and quantization affect the computation of Stokes parameters but many works in the literature, dealing with polarization imaging, simply ignore this aspect. We propose here a simple procedure to improve S_0, S_1, S_2 by assuming that each pixel of $I_0, I_{45}, I_{90}, I_{135}$ is perturbed by zero-mean additive Gaussian noise before being clipped to range $0 \dots 1$.

By observing Eq. 2 we see that our four channel (noisy) polarized image $\bar{I}_{\mathcal{P}} = (\bar{I}_0, \bar{I}_{45}, \bar{I}_{90}, \bar{I}_{135})^T$ is an over-parametrization over the Stokes parameters. Indeed, since:

$$I_0 + I_{90} = I_{45} + I_{135} \tag{11}$$

we can recover any of the four channels by knowing the other three. So, we start by cleaning the acquired images to recover pixels in $\bar{I}_{\mathcal{P}}$ where exactly one of the four channels are either under- or over-exposed.

If all the channels are in the proper range, we consider each pixel p as a vector of four Gaussian distributed random variables $\bar{I}_k(p) = I_k(p) + \epsilon,\ \epsilon \sim N(0, \sigma_{k,p})$ and use a Maximum Likelihood approach to find the best value of $I_k(p)$ such that Eq. 11 holds. In practice this boils down in solving, for each pixel p, the constrained least squares:

$$\begin{aligned} \underset{x_p = \left(I_0(p)\ I_{45}(p)\ I_{90}(p)\ I_{135}(p)\right)^T}{\operatorname{argmin}} \quad & \|x(p) - \bar{I}_{\mathcal{P}}(p)\|^2 \\ \text{s.t.} \quad & \mathbf{C}x = 0 \end{aligned} \tag{12}$$

where $\mathbf{C} = \begin{pmatrix} 1 & -1 & 1 & -1 \end{pmatrix}$. Using the KKT conditions, the analytical optimal solution is given by:

$$\begin{pmatrix} I_0(p) \\ I_{45}(p) \\ I_{90}(p) \\ I_{135}(p) \end{pmatrix} = \begin{pmatrix} k_1 & k_2 & -k_2 & k_2 \\ k_2 & k_1 & k_2 & -k_2 \\ -k_2 & k_2 & k_1 & k_2 \\ k_2 & -k_2 & k_2 & k_1 \end{pmatrix} \begin{pmatrix} \bar{I}_0(p) \\ \bar{I}_{45}(p) \\ \bar{I}_{90}(p) \\ \bar{I}_{135}(p) \end{pmatrix} \tag{13}$$

with coefficients $k_1 = 0.75$ and $k_2 = 0.25$.

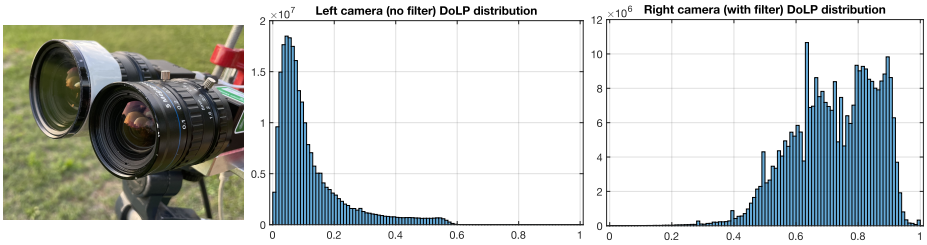

Fig. 3. Left: polarimetric stereo setup (right camera with additional filter on). Center and right: histograms of DoLP distribution for both cameras in our dataset.

3 Experimental Section

In this section we first describe the acquisition process we adopted to compute ground truth images with the proposed setup. Then, we compare the proposed HDR reconstruction technique with state-of-the-art algorithmic and learning-based methods using standard and polarimetric cameras.

An HDR polarimetric dataset was acquired using two FLIR Blackfly monochrome Polarization cameras, mounting a Sony IMX250MZR sensor providing PFA mosaiced images of size 2448×2048 pixels. ICRF was estimated as in [21] and applied to every captured image. The two cameras were mounted on a solid frame with a minimal baseline (see Fig. 3, left), and connected via an hardware trigger so that pictures are acquired simultaneously. As required by our method, an additional external linear polarizer was added to the right camera only, oriented with an angle of $\approx 15°$ wrt. the camera horizontal axis. For each scene, we first took 30 stereo shots with increasing exposure time, then we removed the additional filter and took another set of 30 pictures.

Note that (i) exposure values were set according to the scene and accounting for the external filter attenuation, i.e. the i^{th} intensity image of each set exhibits the same average intensity with and without the filter; (ii) one set of 30 acquisitions takes around half a second and the filter removal is fast ($\sim 1s$). Therefore, the captured scene was essentially unchanged with and without the filter. For each scene, the HDR Ground Truth (GT) was computed by applying [5] to the intensities obtained from the camera with no additional filter. We acquired a total of 20 challenging HDR scenes, both outdoor and indoor.

Table 1. Comparison between HDR methods with Reinhard tone mapping [24] applied.

Method	Best PSNR (dB)	PSNR (dB)	Best MSSIM $\times 10^{-1}$	MSSIM $\times 10^{-1}$
Our	**26.1946 ± 3.5419**	**20.8537 ± 4.8464**	**9.8187 ± 0.1881**	**9.4217 ± 1.0714**
Wu et al. [29]	24.5501 ± 3.1613	19.5490 ± 4.7051	9.7739 ± 0.2428	9.3361 ± 1.1289
HDRCNN	14.3141 ± 1.1944	11.9532 ± 2.3152	7.7141 ± 1.3194	5.9161 ± 2.1367
Deep-HDR	14.8922 ± 1.3342	12.3500 ± 2.5288	7.8611 ± 1.3488	5.9814 ± 2.1720
Two-stage-HDR	19.1025 ± 1.6915	14.4062 ± 3.1722	9.5515 ± 0.2924	8.6506 ± 1.3224
KO	14.1248 ± 1.6047	12.2216 ± 1.8131	5.5661 ± 1.3923	4.4925 ± 1.6370

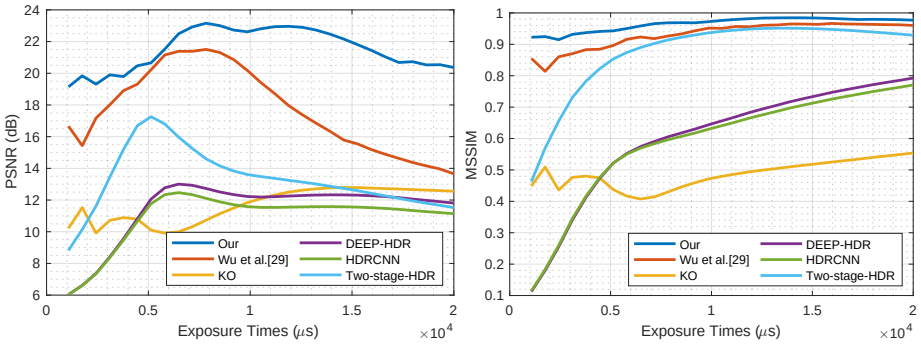

Fig. 4. PSNR (left) and MSSIM (right) for same scene varying exposure times.

3.1 Comparisons

We compared our method with state-of-the-art techniques performing HDR imaging from a single shot. The most similar work proposed in the literature is Wu et al. [29], where HDR is performed using a single PFA camera. We also compared with an inverse tone mapping technique described in [13] (KO from now on) which estimates HDR from a single image. Moreover, we considered three recent learning-based methods namely: HDRCNN [6], Two-stage-HDR [1] and Deep-HDR [26]. These methods are not designed for polarimetric cameras but perform HDR from a single intensity image.

For our method we applied the procedure described in §2.2 (i.e. considering data from both the cameras), while for Wu et al. we only used the polarimetric images captured from the camera with no external filter. Finally, for all the other methods we gave as input demosaiced intensity-only images, simulating a regular grayscale acquisition.

Output HDR images are evaluated against the GT according to their Peak Signal-to-Noise Ratio (PSNR) and Multi-Scale Structural SIMilarity (MSSIM) [28]. Results are listed in Table 1. Best PSNR and best MSSIM columns correspond to average of the best value achieved for each scene among all exposures, while PSNR and MSSIM are average of values computed on all exposures. Our proposed technique produced significantly improved PSNR and MSSIM as compared to other techniques. As expected, Wu et al. exhibits the second best result since it is the only other method taking full advantage of polarimetric information. However, the majority of pixels in the acquired scenes exhibit a DoLP ranging from 0 to 0.2 (See Fig. 3, center) so Wu et.al still suffers a limited extent of equivalent exposure times. On the other hand, adding the external linear polarizer significantly increase the DoLP (Fig. 3, right) thus increasing the dynamic range of the acquisition device.

Since all the approaches use a single-shot to produce the output HDR, it is important to assess how much a correct exposure of the acquired image is critical for the final result. To evaluate this, we plotted in Fig. 4 the resulting PSNR (left) and MSSIM (right) for the same scene varying the exposure time.

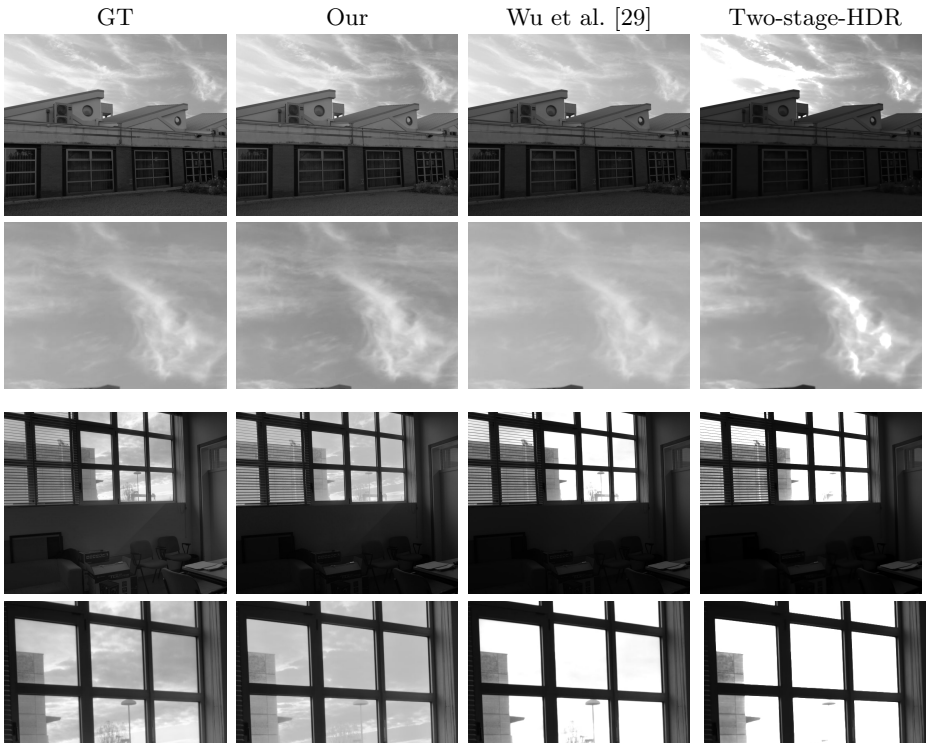

Fig. 5. Qualitative analysis of HDR reconstruction methods. Rows 2 and 4 show a zoomed in patch of scenes shown in 1 and 3 respectively.

All the curves are mostly bell-shaped, showing how there exists a single optimal exposure time maximizing the quality of the resulting HDR image. However, our method is far less sensitive to that (see for example the PSNR plot on the left) where a broad range of exposure times ranging from 6 to 15 ms equally result in a PSNR above 22. In any case, our solution consistently outperforms other techniques demonstrating its versatility in situations where it is not trivial to choose a good exposure time for a certain scene. Finally, Fig. 5 shows some selected examples to qualitatively compare various methods.

4 Conclusions

In this paper we proposed a novel HDR reconstruction technique based on a pair of PFA cameras in which one of the two mounts an additional linear polarizer in front of the lenses. We derived the imaging model for the whole system and showed that the range of equivalent exposure times is much wider than using a single camera as previously proposed by Wu et al. Experiments demonstrate how the new method outperforms state-of-the-art single shot HDR techniques, including recent Deep Learning approaches.

References

1. A Sharif, S., Naqvi, R.A., Biswas, M., Kim, S.: A two-stage deep network for high dynamic range image reconstruction. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 550–559 (2021)
2. Banterle, F., Debattista, K., Artusi, A., Pattanaik, S., Myszkowski, K., Ledda, P., Chalmers, A.: High dynamic range imaging and low dynamic range expansion for generating hdr content. In: Computer graphics forum. vol. 28, pp. 2343–2367. Wiley Online Library (2009)
3. Banterle, F., Ledda, P., Debattista, K., Chalmers, A.: Inverse tone mapping. In: Proceedings of the 4th international conference on Computer graphics and interactive techniques in Australasia and Southeast Asia. pp. 349–356 (2006)
4. Cronin, T.W., Marshall, J.: Patterns and properties of polarized light in air and water. Philosophical Transactions of the Royal Society B: Biological Sciences **366**(1565), 619–626 (2011)
5. Debevec, P.E., Malik, J.: Recovering high dynamic range radiance maps from photographs. In: ACM SIGGRAPH 2008 classes, pp. 1–10 (2008)
6. Eilertsen, G., Kronander, J., Denes, G., Mantiuk, R.K., Unger, J.: Hdr image reconstruction from a single exposure using deep cnns. ACM transactions on graphics (TOG) **36**(6), 1–15 (2017)
7. Endo, Y., Kanamori, Y., Mitani, J.: Deep reverse tone mapping. ACM Transactions on Graphics (Proc. of SIGGRAPH ASIA 2017) **36**(6) (Nov 2017)
8. Farnebäck, G.: Two-frame motion estimation based on polynomial expansion. In: Scandinavian conference on Image analysis. pp. 363–370. Springer (2003)
9. Goldstein, D.H.: Polarized light. CRC press (2017)
10. Kalantari, N.K., Ramamoorthi, R., et al.: Deep high dynamic range imaging of dynamic scenes. ACM Trans. Graph. **36**(4), 144–1 (2017)
11. Khan, E.A., Akyuz, A.O., Reinhard, E.: Ghost removal in high dynamic range images. In: 2006 International Conference on Image Processing. pp. 2005–2008. IEEE (2006)
12. Khan, Z., Khanna, M., Raman, S.: Fhdr: Hdr image reconstruction from a single ldr image using feedback network. In: 2019 IEEE Global Conference on Signal and Information Processing (GlobalSIP). pp. 1–5. IEEE (2019)
13. Kovaleski, R.P., Oliveira, M.M.: High-quality reverse tone mapping for a wide range of exposures. In: 2014 27th SIBGRAPI Conference on Graphics, Patterns and Images. pp. 49–56. IEEE (2014)
14. Kupinski, M.K., Bradley, C.L., Diner, D.J., Xu, F., Chipman, R.A.: Angle of linear polarization images of outdoor scenes. Optical Engineering **58**(8), 082419 (2019)
15. Lee, S., An, G.H., Kang, S.J.: Deep recursive hdri: Inverse tone mapping using generative adversarial networks. In: proceedings of the European Conference on Computer Vision (ECCV). pp. 596–611 (2018)
16. Liu, Y.L., Lai, W.S., Chen, Y.S., Kao, Y.L., Yang, M.H., Chuang, Y.Y., Huang, J.B.: Single-image hdr reconstruction by learning to reverse the camera pipeline. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 1651–1660 (2020)
17. Marnerides, D., Bashford-Rogers, T., Hatchett, J., Debattista, K.: Expandnet: A deep convolutional neural network for high dynamic range expansion from low dynamic range content. In: Computer Graphics Forum. vol. 37, pp. 37–49. Wiley Online Library (2018)

18. McCann, J.J., Rizzi, A.: The art and science of HDR imaging, vol. 26. John Wiley & Sons (2011)
19. Mertens, T., Kautz, J., Van Reeth, F.: Exposure fusion: A simple and practical alternative to high dynamic range photography. In: Computer graphics forum. vol. 28, pp. 161–171. Wiley Online Library (2009)
20. Mihoubi, S., Lapray, P.J., Bigué, L.: Survey of demosaicking methods for polarization filter array images. Sensors **18**(11), 3688 (2018)
21. Mitsunaga, T., Nayar, S.K.: Radiometric self calibration. In: Proceedings. 1999 IEEE computer society conference on computer vision and pattern recognition (Cat. No PR00149). vol. 1, pp. 374–380. IEEE (1999)
22. Pistellato, M., Bergamasco, F., Fatima, T., Torsello, A.: Deep demosaicing for polarimetric filter array cameras. IEEE Transactions on Image Processing **31**, 2017–2026 (2022). https://doi.org/10.1109/TIP.2022.3150296
23. Prabhakar, K.R., Babu, R.V.: Ghosting-free multi-exposure image fusion in gradient domain. In: 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). pp. 1766–1770. IEEE (2016)
24. Reinhard, E., Heidrich, W., Debevec, P., Pattanaik, S., Ward, G., Myszkowski, K.: High dynamic range imaging: acquisition, display, and image-based lighting. Morgan Kaufmann (2010)
25. Santos, M.S., Ren, T.I., Kalantari, N.K.: Single image hdr reconstruction using a cnn with masked features and perceptual loss. arXiv preprint arXiv:2005.07335 (2020)
26. Santos, M.S., Tsang, R., Khademi Kalantari, N.: Single image hdr reconstruction using a cnn with masked features and perceptual loss. ACM Transactions on Graphics **39**(4) (7 2020). https://doi.org/10.1145/3386569.3392403
27. Umow, v.N.: Chromatische depolarisation durch lichtzerstreuung. Phys. Z **6**, 674–676 (1905)
28. Wang, Z., Simoncelli, E.P., Bovik, A.C.: Multiscale structural similarity for image quality assessment. In: The Thrity-Seventh Asilomar Conference on Signals, Systems & Computers, 2003. vol. 2, pp. 1398–1402. Ieee (2003)
29. Wu, X., Zhang, H., Hu, X., Shakeri, M., Fan, C., Ting, J.: Hdr reconstruction based on the polarization camera. IEEE Robotics and Automation Letters **5**(4), 5113–5119 (2020)
30. Yang, X., Xu, K., Song, Y., Zhang, Q., Wei, X., Lau, R.W.: Image correction via deep reciprocating hdr transformation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 1798–1807 (2018)
31. Yu, Y., Zhu, D., Smith, W.A.: Shape-from-polarisation: a nonlinear least squares approach. In: Proceedings of the IEEE International Conference on Computer Vision Workshops. pp. 2969–2976 (2017)
32. Zappa, C.J., Banner, M.L., Schultz, H., Corrada-Emmanuel, A., Wolff, L.B., Yalcin, J.: Retrieval of short ocean wave slope using polarimetric imaging. Measurement Science and Technology **19**(5), 055503 (2008)