

HDR Reconstruction Based On the Polarization Camera

Xuesong Wu¹, Hong Zhang², *Fellow, IEEE*, Xiaoping Hu¹, *Senior Member, IEEE*,
Moein Shakeri², *Member, IEEE*, Chen Fan¹ and Juiwen Ting²

Abstract—The recent development of the on-chip micro-polarizer technology has made it possible to acquire – with the same ease of operation as a conventional camera – spatially aligned and temporally synchronized polarization images simultaneously in four orientations. This development has created new opportunities for interesting applications including those in robotics. In this paper, we investigate the use of this sensor technology in high-dynamic-range (HDR) imaging. Specifically, observing that natural light can be attenuated differently by varying the direction of the polarization filter, we treat the multiple images captured by the polarization camera as a set captured at different exposure times, useful to the reconstruction of an HDR image. In our approach, we first study the radiometric model of the polarization camera, and relate the polarizer direction, degree and angle of polarization of light to the exposure time of a pixel in the polarization images. Subsequently, by applying the standard radiometric calibration procedure of a camera, we recover the camera response function. With multiple polarization images at known pixel-specific exposure times, we can then proceed to estimate the irradiance maps from the images and generate an HDR image. Two datasets are created to evaluate our approach, and experimental results show the dynamic range by our approach can be increased by an amount dependent on light polarization. We also use two robotics experiments on feature matching and visual odometry to demonstrate the potential benefit of this increased dynamic range.

I. INTRODUCTION

The dynamic range of an imaging device is defined by the ratio between the brightest and darkest image irradiance that it is capable of measuring [1]. Conventional imaging technology offers a limited dynamic range, typically eight bits of gray scale or per color channel. This is much narrower than that of the human vision whose dynamic range is on the order of 20 bits [2]. As a result, images of a high-contrast natural scene captured by a camera often contain regions of over- or under-exposed pixels, and fail to convey the true image irradiance in these regions [3]. Considerable research effort has been made to expand the dynamic range of captured images. High-sensitivity imaging sensors can be designed with a higher bit depth than eight. HDR “bracketing” is a technique whereby image irradiance is recovered from multiple images with different exposures [4],

*This work was supported in part by the National Natural Science Foundation of China under Grant 61773394 and Grant 61573371, and by UAHJIC under Grant Number CSI 201902.

¹Xuesong Wu, Xiaoping Hu and Chen Fan are with the College of Intelligence Science and Technology, National University of Defense Technology, Changsha, 410073, China. hellowuxuesong@163.com; xphu.nudt@hotmail.com

²Hong Zhang, Junwen Ting and Moein Shakeri are with the Department of Computing Science, University of Alberta, Edmonton, AB, Canada. hzhang@ualberta.ca

although the scene must remain static with respect to the multiple exposure times to avoid motion blur. Recently, there have also been attempts to estimate an HDR image from a single image by deep neural networks [5], [6]. The learning-based approach offers an additional avenue to pursue to complement the method studied in this paper.

[7] described an interesting sensor design for HDR imaging called spatially varying exposures (SVE) in which pixels are organized into units of four sensing elements each, arranged as a 2x2 matrix, with different sensitivities. SVE equivalently captures four images simultaneously at four different exposure times, t_1 through t_4 , as illustrated in Fig. 1. Once the camera response function is identified, one can estimate the image irradiance and reconstruct an image whose dynamic range is expanded by approximately a factor of four or two bits.

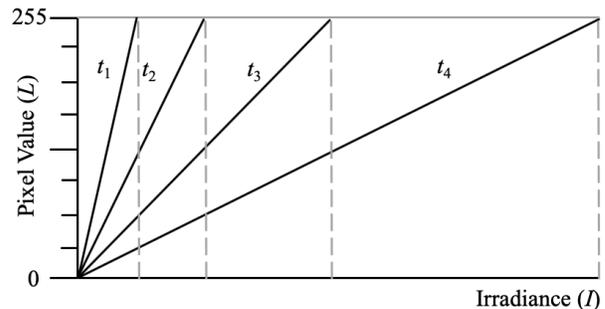


Fig. 1. Illustration of dynamic range expansion with multiple exposures, for a linear camera response function (adopted from [7]). Each exposure time (t_i) maps image irradiance to a pixel value between 0 and 255. In this example, we assume $t_4 < t_3 < t_2 < t_1$.

An interesting recent development in imaging technology is a polarization image sensor (Polarsens by Sony) with four-directional, on-chip micro-polarizers [8]. Similar to SVE, the pixels are organized into computational units. Each unit contains four sensing elements or pixels with polarizing filters at 0° , 45° , 90° , and 135° , respectively, to capture four spatially aligned and temporally synchronized images of a scene (see Fig. 2 for example images). Since polarization conveys information about surface normal, an obvious application of this image sensor is 3D reconstruction [9]. HDR imaging is a less obvious application of this image sensor. Since a polarizing filter attenuates irradiance, and the extent of attenuation varies with the direction of the polarizer, in a way similar to changing the exposure time setting, this sensor also provides us with the possibility to reconstruct HDR images. In fact, Salahieh *et al.* [10] has made such an

attempt based on the idea of light attenuation by a polarizer for the purpose of eliminating saturated points and enhancing contrast to achieve HDR; however, [10] did not examine full HDR image reconstruction.

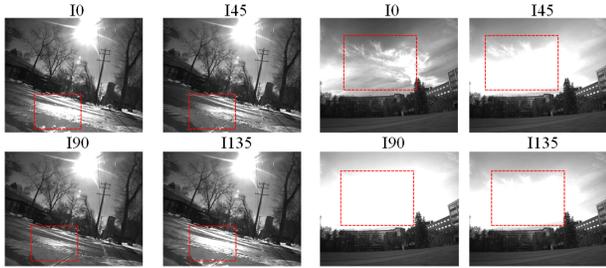


Fig. 2. Sample images of two scenes captured by the polarization camera with the polarizing filters oriented at 0° , 45° , 90° , and 135° respectively. Shown in the red bounding boxes are areas where the pixels are over-exposed or under-exposed in some polarization images but well exposed in others, demonstrating the potential of the polarization camera for HDR image construction.

In this paper, we present our method for using the polarization camera for HDR image reconstruction. We will first present the radiometric model of the polarization camera that maps the image irradiance I_0 to the four polarization images (L_1 through L_4). We use the standard radiometric and polarimetric calibration procedures [1], [11] to identify the response function of the polarization camera. With the help of the radiometric camera model, we conduct an analysis of the expected increase in dynamic range of the reconstructed HDR image from the polarization images. We examine the performance of our proposed method on HDR image reconstruction with the help of datasets we build for outdoor environments. We show experimentally that the our proposed method is capable of increasing the dynamic range of reconstructed images by one to two bits. Finally, we present a couple of examples in robotics that can potentially benefit from the HDR images reconstructed from the polarization camera.

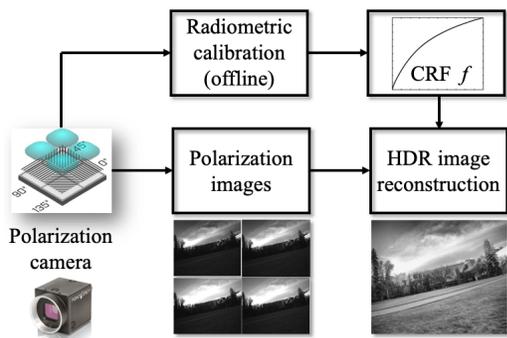


Fig. 3. Overview for the proposed method: images captured by the polarization camera [8] are used to construct an HDR image with the help of the camera response function (CRF), f , identified through offline calibration.

II. PROPOSED HDR RECONSTRUCTION METHOD

In this section, we describe our proposed method for HDR image reconstruction with the polarization camera.

An overview of our method is provided in Fig. 3. We will first present the radiometric model of the polarization camera, which relates image irradiance to the polarization images used in HDR reconstruction. Subsequently, we will conduct an analysis of the change in dynamic range that is potentially possible with the polarization images, based on conditions of light polarization. We will then briefly summarize the calibration process by which to obtain the camera response function. Finally, we provide the solution to HDR reconstruction from polarization images.

A. Radiometric model of the polarization camera

For the sake of introducing terminology, Fig. 4 shows the process in which scene radiance S is perceived by the polarization camera to create four polarization images per exposure. Image irradiance I_0 , an attenuated version of the scene radiance and measured in Jm^{-2} [4], is first filtered by the on-chip directional micro-polarizers. The camera's photo-sensitive elements then convert the light signal into four digital images L_1 through L_4 , through the camera response function. In this process, the camera has an exposure time t_0 (not shown), which can be varied when necessary and affect the polarization images.

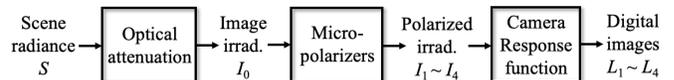


Fig. 4. Image acquisition pipeline shows how scene radiance becomes pixel values for the digital polarization camera.

The effect of a polarizer on image irradiance I_0 can be described by [12]:

$$I_i = \frac{1}{2} \times I_0 [1 + d \cos(2\theta - 2\alpha_i)] \quad (1)$$

where α_i is the angle of the polarizers (at 0° , 45° , 90° and 135° of the polarization camera), and $i = 1, 2, 3, 4$ denotes the index of the four polarizers in a *calculation unit* of the polarization camera. θ ($-90^\circ \leq \theta \leq 90^\circ$) is the angle of polarization (AoP) and d the degree of polarization (DoP). If $d = 0$, the light is unpolarized; if $d = 1$, the light is completely polarized; when $0 < d < 1$, the light is partially polarized. Given the four polarizer directions, the four filtered irradiance maps can be defined by (+ corresponding to I_1 and I_2 , and - to I_3 and I_4 , respectively):

$$\begin{aligned} I_1, I_3 &= I_0(1 \pm d \cos 2\theta) / 2 \\ I_2, I_4 &= I_0(1 \pm d \sin 2\theta) / 2 \end{aligned} \quad (2)$$

From the four pixels within a calculation unit of the polarization camera, it is easy to show that we can estimate d and θ as follows. Defining the Stokes parameters as

$$S_0 = \frac{1}{2} (I_1 + I_2 + I_3 + I_4), \quad S_1 = I_1 - I_3, \quad S_2 = I_2 - I_4, \quad (3)$$

d and θ are computed from the Stokes by [12]

$$d = \frac{\sqrt{S_1^2 + S_2^2}}{S_0}, \quad \theta = \frac{1}{2} \tan^{-1} \left(\frac{S_2}{S_1} \right) \quad (4)$$

Note that d and θ vary spatially with pixel locations.

In general, the relationship between polarized irradiance I_i and perceived pixel value L_i by the sensor at exposure time t_0 is defined by the camera response function [4]:

$$L_i = f(I_i t_0) \quad \text{where } i = 1, 2, 3, 4 \quad (5)$$

Substituting Eq. (2) into Eq. (5), we obtain

$$L_i = f(I_0 t_i) \quad (6)$$

the so-called reciprocity equation [4] where

$$t_1, t_3 = \frac{t_0}{2}(1 \pm d \cos 2\theta), \quad \text{and } t_2, t_4 = \frac{t_0}{2}(1 \pm d \sin 2\theta). \quad (7)$$

Eq. (7) shows that when incoming light is not entirely unpolarized ($d \neq 0$), the four pixels within one calculation unit of the polarization camera experience different exposure times, effectively creating the condition for multiple exposures. Different from a conventional camera undergoing multiple exposures, the variation in exposure time in the polarization camera is pixel specific as both d and θ vary from pixel to pixel. This is because the pixels in different calculation units of the polarization camera correspond to different points in space, which in general differ in terms of their light polarization, just as in color or in intensity.

B. Dynamic range of the polarization camera

In general, the dynamic range of an image sensor is defined as:

$$DR = 20 \log \left(\frac{C_{full}}{N_r} \right) \quad (8)$$

where C_{full} represents the full-well capacity of the sensor and N_r the read-noise [7]. Typically, the maximum gray level L_{max} corresponds to the full-well capacity and the minimum level L_{min} corresponds to the minimum signal detectable by the sensor. Computationally, for an ideal 8-bit digital camera with a linear response function, its dynamic range

$$DR = 20 \log \left(\frac{L_{max}}{L_{min}} \right) = 20 \log(255) = 48.13 \text{ dB}.$$

The sensed dynamic range can be increased by multiple exposures. The expanded dynamic range is related to the ratio between the maximum exposure time t_{max} and the minimum exposure time t_{min} by [7]:

$$DR = 20 \log \left(\frac{L_{max} t_{max}}{L_{min} t_{min}} \right) \quad (9)$$

For the polarization camera, $t_{max} = \max(t_1, t_2, t_3, t_4)$ and $t_{min} = \min(t_1, t_2, t_3, t_4)$ within a calculation unit. The dynamic range of the entire camera can be readily calculated accordingly, using the mean of the dynamic ranges of all its calculation units.

To quantify the gain in dynamic range of the images that can be potentially obtained, Table I summarizes the results based on the evaluation of Eq. (9) under three different values of d . In the calculation, we assume a uniform distribution of

the angle of polarization. We also calculate the change in dynamic range in terms of the number of additional bits in the pixel depth, i.e., $\log_2 \left(\frac{t_{max}}{t_{min}} \right)$. At $d = 0.2$ or a moderate amount of polarization in the environment light, we can expect to increase the dynamic range by 0.53 bits or 3.2 dB. On the other hand, for space points with significant light polarization with $d = 0.8$, the dynamic range of the images can be increased by 2.66 bits, from 8 bits to 10.66 bits, or by 16.4 dB, from 48.13 dB to 64.53 dB. In Table I, the standard deviation of the change in dynamic range (Δ) is due to the variation in the angle of polarization (θ). Practically, d can vary spatially significantly and, as a result, so can the different regions of an image in terms of their gain in dynamic range from this polarization camera.

TABLE I
CHANGE IN DYNAMIC RANGE \pm ONE STANDARD DEVIATION AS A
FUNCTION OF DEGREE OF POLARIZATION (d)

d		0.2	0.5	0.8
Δ	DR (dB)	3.2	8.5	16.4
	# of bits	0.53 \pm 0.05	1.40 \pm 0.16	2.66 \pm 0.41

C. Radiometric calibration of the polarization camera

As shown in Fig. 3, HDR reconstruction in general requires the knowledge of the camera response function, which maps image irradiance to image brightness or pixel value. Rich literature exists on methods that allow one to calibrate a camera and obtain its response function. Most consumer cameras have a nonlinear response function. In our study, we find experimentally that the method of Mitsunaga and Nayar [1] is able to calculate the inverse of the CRF accurately, among other methods. Most importantly, it is formulated in such a way as is applicable to the polarization camera, as will be explained shortly. We briefly summarize the method in this section for the completeness of the presentation.

Define the camera response function as

$$L = f(I t)$$

where I , L and t are the image irradiance, the pixel value and the exposure time, respectively. Rather than identifying f , [1] estimates its inverse or the inverse camera response function (ICRF), g , defined by

$$I = \frac{g(L)}{t} \quad (10)$$

which is of direct use in HDR reconstruction. g can be modeled using a polynomial of appropriate order N :

$$g(L) = \sum_{n=0}^N c_n L^n \quad (11)$$

With this formulation, the calibration process is viewed as one of determining the order N as well as the coefficients c_n . When multiple images are taken at different exposures (aperture multiplied by shutter time), the calibration algorithm makes use of the observation that the ratio between

exposures is the same as that between the (scaled) scene radiance as is defined by Eq. (11). Using the multiple available ratios between known exposures, this observation is used to set up constraints on the coefficients c_n of Eq. (11) whose pixel value L is measured and therefore known. For the polarization camera, because the basic constraint used in [1] takes the form of a ratio between scene radiance, the filtering effect of the polarizer appears on both numerator and the denominator of the ratio and cancels each other. The algorithm can therefore be used as is to calibrate the polarization camera.

We have used the above calibration method to perform radiometric calibration of our polarization camera, and the ICRF is shown in Figure 5. Note that both the pixel intensity measurement and the scene radiance are on a normalized scale. In our experiment, we used a total of seven different exposures (in the UAPolCampus dataset, to be detailed in Section III.B). We initially set N to 6, and stepped down incrementally for optimal fit with the image data. We found $N = 2$ to provide the optimal result, with $c_0 = 0.002$, $c_1 = 0.5278$, and $c_2 = 0.4702$; goodness of fit in terms of R-squared is 0.962. This ICRF is then used to reconstruct HDR images, as will be described in the next section.

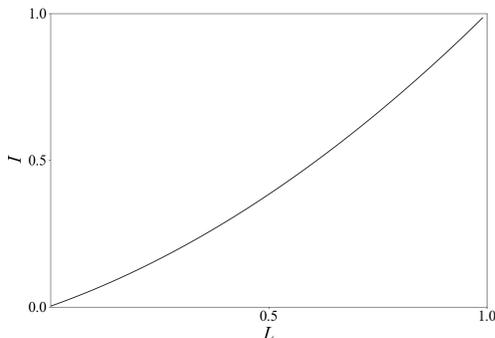


Fig. 5. Inverse camera response function of the polarization camera found by the radiometric calibration algorithm in [1].

D. Reconstructing the high dynamic range images

Once the inverse camera response function g is identified, it can be used to convert pixel values to relative irradiance values. We use the inverse CRF defined by Eq. (10) four times on the four pixels within a calculation unit described by Eq.(6) with their exposure times t_i , to obtain:

$$I_0 = \frac{g(L_i)}{t_i} \text{ for } i = 1, 2, 3, 4 \quad (12)$$

To reconstruct a pixel of the HDR image, we computed the weighted sum of the irradiance values of the four pixels in one calculation unit where the weights depend on how well the pixels are exposed [4]:

$$I_0 = \frac{\sum_{i=1}^4 W(L_i)(g(L_i)/t_i)}{\sum_{i=1}^4 W(L_i)} \quad (13)$$

W is the weight function, which can be implemented with the Gaussian function [13] (with a standard deviation of $\sigma = 0.2$ in our study): $W(L_i) \sim \exp(-\frac{(L_i-0.5)^2}{2\sigma^2})$.

In practice, the equivalent exposure time t_i is essential for the reconstruction of the HDR image according to Eq. (13), and it is calculated with Eq. (7) from d (DoP) and θ (AoP). However, since d and θ are estimated from the four pixels in each calculation unit [14], for pixels that are under- or over-exposed, d and θ cannot be calculated accurately. To solve this problem, we first substitute Eq. (7) into Eq. (10) to obtain

$$\begin{aligned} g(L_1), g(L_3) &= \frac{I_0 t_0}{2} (1 \pm d \cos 2\theta) \\ g(L_2), g(L_4) &= \frac{I_0 t_0}{2} (1 \pm d \sin 2\theta) \end{aligned} \quad (14)$$

and then manipulate Eq. (14) to arrive at two expressions for I_0 that are independent of the exposure times t_i :

$$I_0 = \frac{g(L_1) + g(L_3)}{t_0}, \quad I_0 = \frac{g(L_2) + g(L_4)}{t_0} \quad (15)$$

Although either of the above two equations can be used alone to estimate I_0 , their weighted sum is used instead in the same way as Eq. (13) for robustness:

$$I_0 = \frac{\sum_{i=1}^2 W(L_i + L_{i+2})(g(L_i) + g(L_{i+2}))/t_0}{\sum_{i=1}^2 W(L_i + L_{i+2})} \quad (16)$$

The sums $L_i + L_{i+2}$ are normalized to $[0,1]$ before they are used to compute the weights.

Algorithm 1: HDR Image Reconstruction

Input: L_1 thru L_4 , t_0 and ICRF g

Output: HDR I_0

- 1 Construct polarized irradiance maps $I_i = g(L_i)/t_0$;
 - 2 Calculate d and θ with Eqs. (3) and (4);
 - 3 **for each calculation unit do**
 - 4 Calculate t_1 through t_4 with Eq. (7);
 - 5 Estimate an I_0 from each of 4 pixels by Eq. (12);
 - 6 Combine the multiple I_0 with Eq. (13) or (16);
 - 7 **end**
-

Note that Line 1 in Algorithm 1 comes from Eqs. (5) and (10).

III. EXPERIMENTAL EVALUATION

In this section, we present the experimental results that verify our proposed method for HDR image reconstruction and study its use in a couple of example robotics applications. We will first validate the radiometric model of the polarization camera developed in Section II.A, particularly with regard to whether the camera is capable of capturing synchronized images at multiple exposures. We will then proceed to introduce two datasets that we have constructed to study the HDR reconstruction algorithm in Section II.D. We compare our algorithm with standard inverse tone mapping methods for HDR reconstruction, to show the importance of our radiometric camera model. Finally, we will examine

if the HDR images from our reconstruction algorithm can improve the performance of two robotics applications: image matching and visual odometry.

A. Validation of the radiometric camera model

The radiometric camera model is verified with the help of a liquid crystal display monitor, which is simply the screen of a laptop that generates pure polarized light [11]; as a result, d and θ are constant within an experiment. For properly exposed pixels with values in the interior of the measurement range, we can reasonably consider the CRF as linear with the pixel value changing linearly with the exposure time. We set t_0 to various exposure values and, for each, we use Eq. (7) to predict t_1, t_2, t_3 and t_4 – the four exposure times within a calculation unit of the polarization camera. We also record their pixel values. We subsequently set the exposure time of the camera t_0 to the predicted t_1 (or any of t_2, t_3 and t_4), and add L_1 and L_3 (or equivalently L_2 and L_4). Under the linear CRF assumption, according to our radiometric model Eq. (2), the summed pixel value should be close to the the pixel value L_1 recorded at the previous exposure time t_0 . We collect the results on 11 randomly chosen calculation units for 44 exposure times t_0 . The error between pixel values is 3.5 on a scale of 256 or less than 3% on average with a standard deviation of 3.36, confirming the accuracy of Eq. (7).

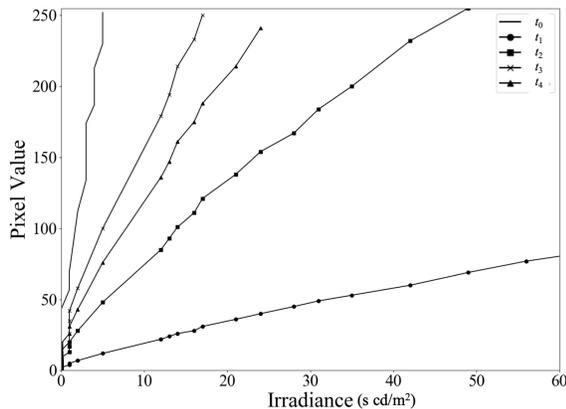


Fig. 6. Measured camera responses based on controlled polarized lighting.

To further verify the radiometric model of our polarization camera, we create controlled polarized light with a dome light source, a photometer and a linear polarizer. The dome light is passed through the linear polarizer to create highly polarized light at $d = 0.85$ and $\theta = -82^\circ$. With the photometer, we measure the radiance of the polarized light, and its range lies between 0 and 70135 cd/m^2 . In combination, we can also vary the exposures of the camera from 0 to $3.86 \times 10^2 \text{ cd} \times \text{s} \times \text{m}^{-2}$. We record the pixel intensities of the polarization camera as well as that of a regular camera. The five resulting camera response curves (one from the regular camera and four from the polarization camera) are shown in Fig. 6. The slopes of the CRF for the polarization camera are dependent on the directions of the four polarizers as defined by Eq. (7). This set of curves match the expected behavior of multiple exposures created by the polarization

camera in Fig. 1. The Pearson correlation coefficient between the predicted slopes of the linear camera response function according Eq.(7) and the slopes of the measured data is 0.997, indicating strong agreement.

B. HDR image reconstruction

We have also conducted experiments to study our proposed HDR image reconstruction algorithm. Due to the lack of publicly available datasets of polarization images at multiple exposure times, we build two datasets, namely *UAPolCampus dataset* and *EdPolCommunity dataset*. Each dataset consists of grayscale images of four channels in 30 different scenes with multiple exposure times in outdoor environments. Specifically, the seven exposure times used in creating *UAPolCampus* are:

$$t_0 = [0.066, 0.125, 0.25, 0.5, 1, 2, 4];$$

and the eight in creating *EdPolCommunity* are:

$$t_0 = [0.1, 0.2, 0.5, 1, 2, 4, 8, 16]$$

all in milliseconds. Exposing a scene multiple times allows us to construct an HDR image with a traditional technique to be compared with our HDR reconstruction algorithm. Sample images of the *EdPolCommunity* dataset are shown in Fig. 8. Each polarization image is of 1224×1024 pixels.

The distributions of d in the two datasets, calculated from the pixels of all their images, are similar, and that of the *UAPolCampus* dataset is plotted in Fig. 7, with an average of 0.1. As can be seen, most pixels have a low d , although the distribution has a long tail. The small peak around $d = 0.4$ indicates significant polarization for some pixels, which can be exploited by the proposed HDR reconstruction method. In terms of the dynamic range, the average DR of all the calculation units is 49.83 dB or 8.28 bits in pixel depth, and the maximum DR is 68.95 dB or 11.45 bits.

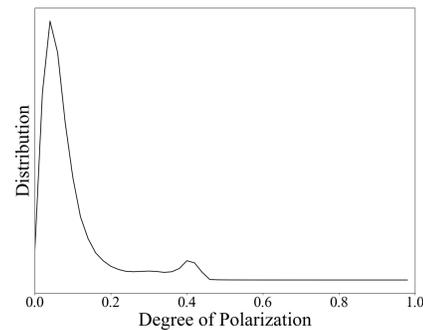


Fig. 7. (a) The distribution of the Degree of Polarization in Dataset *EdPolCommunity*. The average DoP is 0.1.

With the multiple exposures, we create the “ground-truth” HDR images by first creating two HDR polarization images from the two channels of the polarization camera (e.g., L_1 and L_3). For this, we use the method described in [3] and publicly available in MATLAB. The two HDR polarization images are then added to create the final ground-truth HDR

image. The addition removes the effect of the polarizers as is shown by Eq. (15).



Fig. 8. EdPolCommunity dataset with 30 different scenes.

We compare our HDR reconstruction method with several well-known inverse tone mapping (ITM) methods on dataset EdPolCampus. ITM estimates an HDR image from one or several LDR images captured at multiple exposures by first estimating the CRF and then the irradiance map. We compare our method with the competing ITM methods run on a single image, with sample results shown in Figure 9. Visually, the proposed method is able to preserve the details both on the lawn and in the sky with higher contrast better than the competing ITM methods. To perform quantitative experiments, we use the HDR Toolbox [3], which is a MATLAB-based library for processing HDR content. For the parameter settings, we use $\gamma = 1/0.8 = 1.25$ for inverse gamma correction in all existing methods.

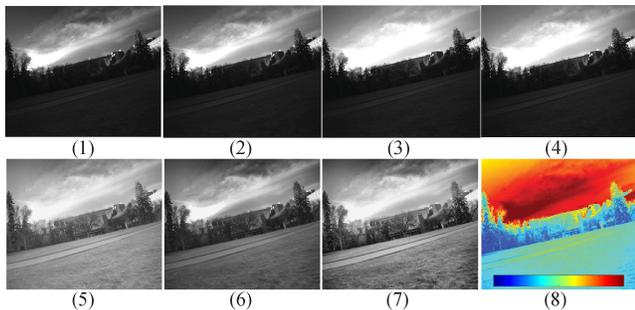


Fig. 9. Raw polarization images(1)(2)(3)(4), tonemapped results of inverse tone mapping method [15] (5), [16] (6) and our proposed method (7). (8) is the false color image of our result to visually show the HDR irradiance map from the low of $5.4e+00$ lux (blue) to the high of $3.9e+02$ lux (red).

With the ground truth HDR images being the reference output, we use PSNR and SSIM [17] as performance metrics to evaluate the competing methods. Outputs of all methods are tone-mapped by Reinhard [18] to LDR for calculating the metrics. As a performance metric, we use HDR-VDP-2 [19], which is based on the human visual system. Table II summarizes the comparison of the proposed method and the optimal result of the ITM methods among the four channels

at one exposure time. We can see that the proposed method is superior to the other methods in all the metrics.

TABLE II
COMPARISON BETWEEN THE EXISTING INVERSE TONE MAPPING METHODS [3] AND OUR PROPOSED HDR RECONSTRUCTION METHOD, WITH RESPECT TO TWO REFERENCE METRICS (PSNR AND SSIM) AND ONE NON-REFERENCE METRIC (HDR-VDP2). THE NUMBERS ARE EXPRESSED IN TERMS OF MEAN \pm STANDARD DEVIATION.

Method	PSNR	SSIM	HDR-VDP2 Q-score
AO	41.2 \pm 10.1	0.92 \pm 0.05	42.5 \pm 4.0
HEO	37.9 \pm 10.3	0.85 \pm 0.07	23.6 \pm 4.3
HPEO	29.6 \pm 8.4	0.90 \pm 0.05	43.5 \pm 3.9
KEO	39.5 \pm 12.0	0.91 \pm 0.05	42.2 \pm 4.0
KOEO	41.3 \pm 10.2	0.92 \pm 0.05	42.8 \pm 4.0
LEO	37.4 \pm 9.8	0.90 \pm 0.05	25.9 \pm 1.5
MEO	35.2 \pm 12.4	0.91 \pm 0.05	42.5 \pm 3.8
Ours	44.5 \pm 10.0	0.96 \pm 0.04	45.6 \pm 5.0

C. Robotics Applications

We have also conducted experiments to determine if the proposed HDR reconstruction method is able to benefit robotics applications. Specifically, we examine feature detection and visual odometry as examples. Previous research has shown that HDR images can improve feature detection [20]. We are interested in determining if the same conclusion can be reached with our HDR images from the polarization camera. In both examples, our results are preliminary and empirical with HDR images showing improvement.

Our first experiment involves feature detection in 30 pairs of images of the UAPolCampus dataset at two exposure times of 1 ms and 0.2 ms. We use L_2 to represent LDR images in this experiment, and use our proposed method to create the HDR images. Both the LDR and the HDR images are tone-mapped to 8-bits by the CLAHE method [21] to suit the subsequent feature detectors. SURF [22], Harris corner [23] and FAST [24] detectors are chosen in our experiment. The repeatability score [25] is computed for the three keypoint detectors. The results of the comparison between LDR and HDR are shown in Table III where we observe a slight but definitive improvement in all three detectors. This conclusion is consistent with that in [20].

TABLE III
COMPARISON BETWEEN LDR AND HDR IMAGES IN TERMS OF REPEATABILITY OF THREE KEYPOINT DETECTORS

Tone-Mapped	Repeatability		
	SURF	Harris	FAST
LDR images	0.784	0.506	0.940
HDR images	0.815	0.640	0.948

In the second experiment, we run DSO (direct sparse odometry) [26] in an indoor environment with two flat walls to estimate its 3D geometry. DSO reconstructs 3D points from pixels of high gradient magnitude by matching image

patches and could potentially benefit from improved image quality. Fig 10 shows the experimental results where the depth values of the 3D points are color-coded [26]. For points on the first wall in red ellipses, DSO with HDR images is able to reconstruct with a higher accuracy than that with regular images as can be judged by the color consistency of the pixels in the enclosed areas. For the second wall, however, HDR images offer a comparable performance to regular images. Quantitatively, for the first wall, DSO with HDR images is able to correctly reconstruct about 10% more 3D points than with regular images.

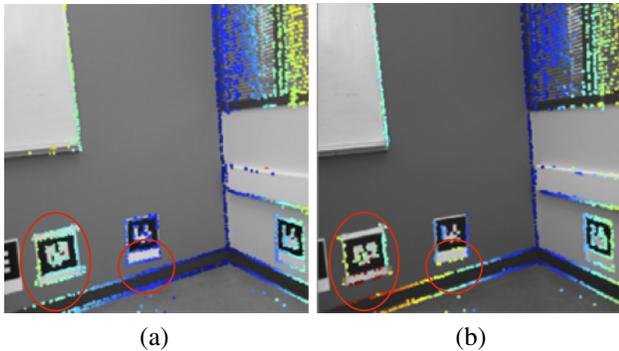


Fig. 10. Comparison of 3D reconstruction by DSO from (a) HDR images obtained with the proposed method and (b) regular images. The depth value is color-coded [26]. Points in the red ellipses on the first wall clearly show more accurate 3D reconstructed with HDR images than with regular images, as judged the consistency of the pixel colors; however, reconstruction quality is similar for the second wall between the two types of images.

IV. CONCLUSION

In this paper, the feasibility of creating high dynamic range images from a camera with on-chip multi-directional polarization filters is studied. By exploiting the fact that in environments with polarized lighting, the effect of the polarizers is analogous to that of imaging with multiple exposures. This gives rise to the possibility of reconstructing the HDR image of an environment from the polarization images. We have developed the radiometric model of the polarization camera with which we can estimate the expected increase in dynamic range as a function of the degree of polarization of the light. We have verified this model using a controlled polarized lighting source and with two datasets we have built for this research.

Two types of experiments have been conducted to demonstrate the benefit of the proposed HDR imaging method. First, we have compared our method with existing inverse tone mapping methods to show the importance of considering the spatial variation of the exposure time when using the polarization camera to estimate the HDR image irradiance. Secondly, we have provided empirical evidence that, in the robotics applications of feature detection and visual odometry, the polarization camera could potentially improve the performance. We leave the systematic investigation of the proposed HDR imaging method in robotics applications as our future work.

REFERENCES

- [1] T. Mitsunaga and S. K. Nayar, "Radiometric self calibration," in *Proceedings. 1999 IEEE CVPR (Cat. No PR00149)*, vol. 1. IEEE, 1999, pp. 374–380.
- [2] K.-F. Yang, X.-S. Zhang, and Y.-J. Li, "A biological vision inspired framework for image enhancement in poor visibility conditions," *IEEE Transactions on Image Processing*, vol. 29, pp. 1493–1506, 2020.
- [3] F. Banterle, A. Artusi, K. Debattista, and A. Chalmers, *Advanced high dynamic range imaging*. CRC press, 2017.
- [4] P. E. Debevec and J. Malik, "Recovering high dynamic range radiance maps from photographs," in *ACM SIGGRAPH 2008 classes*, pp. 1–10.
- [5] R. Gomez-Ojeda, Z. Zhang, J. Gonzalez-Jimenez, and D. Scaramuzza, "Learning-based image enhancement for visual odometry in challenging hdr environments," in *2018 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2018, pp. 805–811.
- [6] X. Yang, K. Xu, Y. Song, Q. Zhang, X. Wei, and R. W. Lau, "Image correction via deep reciprocating hdr transformation," in *Proceedings of the IEEE CVPR*, 2018, pp. 1798–1807.
- [7] S. K. Nayar and T. Mitsunaga, "High dynamic range imaging: Spatially varying pixel exposures," in *Proceedings IEEE CVPR 2000 (Cat. No. PR00662)*, vol. 1. IEEE, 2000, pp. 472–479.
- [8] "Polarization image sensor," <https://www.sony-semicon.co.jp/e/products/IS/polarization/product.html>.
- [9] Z. Cui, J. Gu, B. Shi, P. Tan, and J. Kautz, "Polarimetric multi-view stereo," in *Proceedings of the IEEE CVPR*, 2017, pp. 1558–1567.
- [10] B. Salahieh, Z. Chen, J. J. Rodriguez, and R. Liang, "Multi-polarization fringe projection imaging for high dynamic range objects," *Optics express*, vol. 22, no. 8, pp. 10 064–10 071, 2014.
- [11] Z. Wang, Y. Zheng, and Y.-Y. Chuang, "Polarimetric camera calibration using an lcd monitor," in *Proceedings of the IEEE CVPR*, 2019, pp. 3743–3752.
- [12] D. H. Goldstein, *Polarized light*. CRC press, 2017.
- [13] 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. Wiley Online Library, 2009, pp. 161–171.
- [14] G. Han, X. Hu, J. Lian, X. He, L. Zhang, Y. Wang, and F. Dong, "Design and calibration of a novel bio-inspired pixelated polarized light compass," *Sensors*, vol. 17, no. 11, p. 2623, 2017.
- [15] R. P. Kovaleski and M. M. Oliveira, "High-quality reverse tone mapping for a wide range of exposures," in *2014 27th SIBGRAPI Conference on Graphics, Patterns and Images*, 2014, pp. 49–56.
- [16] H. Landis, "Production-ready global illumination," *Siggraph course notes*, vol. 16, no. 2002, p. 11, 2002.
- [17] Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multiscale structural similarity for image quality assessment," in *The Thirty-Seventh Asilomar Conference on Signals, Systems & Computers*, 2003, pp. 1398–1402.
- [18] 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.
- [19] R. Mantiuk, K. J. Kim, A. G. Rempel, and W. Heidrich, "Hdr-vdp-2: A calibrated visual metric for visibility and quality predictions in all luminance conditions," *ACM Transactions on graphics (TOG)*, vol. 30, no. 4, pp. 1–14, 2011.
- [20] A. Rana, G. Valenzise, and F. Dufaux, "Evaluation of feature detection in hdr based imaging under changes in illumination conditions," in *2015 IEEE ISM*. IEEE, 2015, pp. 289–294.
- [21] K. Zuiderveld, "Contrast limited adaptive histogram equalization," in *Graphics gems IV*, 1994, pp. 474–485.
- [22] H. Bay, T. Tuytelaars, and L. Van Gool, "Surf: Speeded up robust features," in *ECCV*. Springer, 2006, pp. 404–417.
- [23] C. G. Harris *et al.*, "A combined corner and edge detector," in *Alvey vision conference*, vol. 15, no. 50. Citeseer, 1988, pp. 10–5244.
- [24] E. Rosten and T. Drummond, "Machine learning for high-speed corner detection," in *ECCV*. Springer, 2006, pp. 430–443.
- [25] K. Mikolajczyk, T. Tuytelaars, C. Schmid, A. Zisserman, J. Matas, F. Schaffalitzky, T. Kadir, and L. Van Gool, "A comparison of affine region detectors," *IJCV*, vol. 65, no. 1-2, pp. 43–72, 2005.
- [26] J. Engel, V. Koltun, and D. Cremers, "Direct sparse odometry," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Mar. 2018.