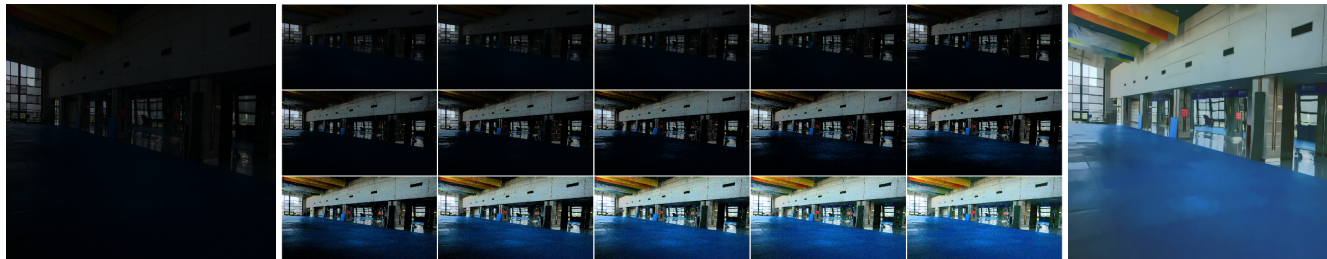# Quaternion Factorized Simulated Exposure Fusion

Saurabh Saini
CVIT, IIIT-Hyderabad
Hyderabad, India
saurabh.saini@research.iiit.ac.in

P J Narayanan
CVIT, IIIT-Hyderabad
Hyderabad, India
pjn@iiit.ac.in

Figure 1: QFSEF: Given a poorly lit image as input (left), we factorize it into multiple illumination consistent layers using a pure quaternion matrix factorization scheme (section 3), which we then use to simulate an exposure stack (mid) and fuse to obtain an enhanced image (right).

ABSTRACT

Image Fusion maximizes the visual information at each pixel location by merging content from multiple images in order to produce an enhanced image. Exposure Fusion, specifically, fuses a bracketed exposure stack of poorly lit images to generate a properly illuminated image. Given a single input image, exposure fusion can still be employed on a 'simulated' exposure stack, leading to direct single image contrast and low-light enhancement. In this work, we present a novel 'Quaternion Factorized Simulated Exposure Fusion' (QFSEF) method by factorizing an input image into multiple illumination consistent layers. To this end, we use an iterative sparse matrix factorization scheme by representing the image as a two-dimensional pure quaternion matrix. Theoretically, our representation is based on the dichromatic reflection model and accounts for the two scene illumination characteristics by factorizing each progressively generated image into separate specular and diffuse components. We empirically prove the advantages of our factorization scheme over other exposure simulation methods by using it for the low-light image enhancement task. Furthermore, we provide three exposure fusion strategies which can be used with our simulated stack and provide a comprehensive performance analysis. Finally, in order to validate our claims, we show extensive qualitative and quantitative comparisons against relevant state-of-the-art solutions on multiple standard datasets along with relevant ablation analysis to support our proposition. Our code and data are publicly available for easy reproducibility and reference. [1]

[1] https://github.com/sophont01/QFSEF

CCS CONCEPTS

• **Computing methodologies** → *Image representations*; **Image processing**; *Optimization algorithms*; **Computational photography**.

KEYWORDS

Low Light Image Enhancement, Quaternion Image Representation, Exposure Fusion, Robust Principal Component Analysis

1 INTRODUCTION

Image enhancement is a classic Computer Vision application involving problems like image denoising, super resolution, contrast enhancement, deblurring *etc*. Low light image enhancement is also a well studied research problem with applications benefiting image classification, recognition, reconstruction *etc*. Several solutions have been proposed to this end using a variety of inputs like flash-noflash image pairs, bracketed exposure sequences, light field data, raw sensor signals *etc*. Flash-noflash images enable a range of applications like material estimation, light spectra estimation, illumination map estimation *etc*. but the with the core limitation that scene areas far away from the flash illumination are not resolved properly in the result. On the other hand models which assume raw images from the camera or light field data as inputs are restrictive in everyday applications. Comparatively it is easy to obtain a bracketed sequence of images with varying shutter speed (*i.e.* an exposure stack) using already available camera modes and external installable softwares. We focus on this category of low-light image enhancement methods and propose a technique to simulate an exposure stack from a single poorly illuminated input image.
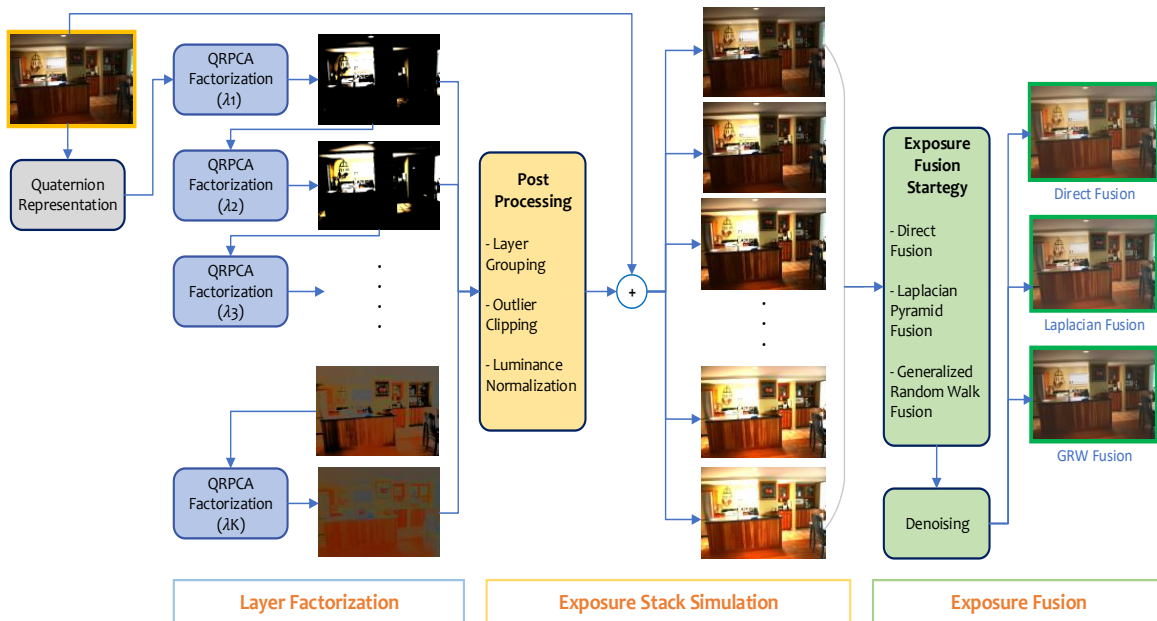
Figure 2: Conceptual Overview: An abstract block diagram of our proposed approach shows our three system sub-modules: *Layer Factorization*, applies gradually relaxed iterative RPCA on Quaternion representation of the input image, followed by *Exposure Stack Simulation*, where we combine factors with original image to render controlled illumination image sequence and finally *Exposure Fusion* where we merge the stack information via three strategies to obtain enhanced results.

Several stacked image representations have been used in the literature for multiple applications. Apart from the exposure stacks, bracketed sequences of images with varying depth-of-field have also been used in several papers [30]. Several methods also exist which use layered depth information for view synthesis and scene understanding [53]. Semantic information has also been factorized into multiple layers using soft segmentation, color unmixing and image matting in [4, 5]. The main advantage of layered representations is that the similar image regions are grouped together which allows easy global manipulation without requiring adaptive local adjustments. Additionally, compared to patch based processing, our estimated layers have an intuitive semantic meaning and can be directly utilized by end users for preferential adjustments.

Single image low-light enhancement is a challenging task considering the limited amount of information available. This is further exacerbated due to camera sensor noise prevalent in the dark image regions, which, if not handled properly will lead to colored artifacts in the results. Earlier solutions for this task were mostly optimization based involving histogram equalization, tone curve adjustment, Retinex theory based illumination map estimation *etc.* Recent advances in the field mostly employ deep learning techniques and train priors using supervision on large datasets. Current state-of-the-art methods are inspired by the fundamentals of the the problem and try to encode domain expertise in their systems *e.g.* decomposing the problem using Retinex theory [56], constructing multiscale Laplacian pyramids [2], adaptable tonal curve adjustment [25] *etc.*

In this work, we present a new insight into the problem by proposing to factorize an image into a sequence of specular factors. We propose a method to progressively remove the specular content from the image by performing iterative sparse matrix factorization which can be used to render a virtual exposure stack. This converts the difficult single image exposure correction problem into a simpler exposure fusion task thereby allowing us to directly use existing exposure fusion strategies on a single image. In order to make this simple idea work, we harness the power of *Quaternions* by representing the image pixel color values on a unit norm sphere of pure quaternions (*i.e.* no real component). This enables better intercolor spectral representation across channels compared to the traditional concatenated color channel approach and allows for a simple yet elegant and interpretable solution.

To summarize, following are our main contributions in this work:

○ We present a novel single image exposure fusion method by simulating an exposure stack for low light image enhancement application.
○ We present an iterative sparse matrix factorization scheme for specular content estimation by using Robust Principal Component Analysis in quaternion space.
○ We provide performance scores and comparison with the current state-of-the-art solutions via extensive experiments and ablation analyses.

2 RELATED WORK

Quaternion Image Processing: Apart from the wider utility of quaternions in robotics and graphics for 3D manipulation, they have also been used in several image processing tasks. Quaternion counterparts to several crucial techniques like Fourier transforms [19], Singular Value Decomposition (SVD) [48], Principal Component Analysis (PCA) [9], derivatives like Gradients and Hessians [57], and optimizations like least square algorithm [34] *etc.* have been proposed in the literature and newer techniques are being developed. Following this, several methods based on the application of these techniques have also come up. Initially Sangwine [47] used quaternion image representation for the task of color image edge detection and proposed additional quaternion filters in [45]. Later in [19] authors formulated vector Fourier transform for quaternions which was then used for saliency prediction [49], texture estimation [7], motion detection [6], image smoothing [51] *etc.* Recently they have also been used for face recognition [33], image inpainting [31] and in image forensics [59]. Here, we use one such existing technique and propose a novel low-light illumination solution. Specifically we use the quaternion RPCA formulation proposed by Chan and Yang [16] who employ it for the task of separating relatively sparse human voice signals from the background musical score in a short music clip.

Robust Principal Component Analysis: RPCA as used in this work, was first proposed by Candès et al. [15] in which they presented a closed form solution to the sparse *vs.* low-rank matrix decomposition problem using a convex optimization algorithm named Principal Component Pursuit (PCP). The method is named RPCA because it is able to recover matrix principal components even in the case of corrupted or missing values. Candès et al. [15] illustrated the utility of their algorithm by using it in two computer vision tasks: video background removal and specularity estimation. Since its initial proposition, RPCA problem has been solved using a variety of optimization methods [11] and used in a variety of applications [10] in several domains like low level image analysis, medical imaging, 3D computer vision and video processing [17, 24, 29, 54, 63]. We take inspiration from the initial application of RPCA by its authors and use it for image specularity estimation.

Illumination Analysis: From image intrinsics perspective (*i.e.* Retinex theory) [8, 35], there are two fundamental components in the image formation process *i.e.* object material dependent reflectance and scene illumination dependent shading [42–44]. Either or both can be used for the task of image based rendering which involves generating a new image directly from a given input image. Controlling the first reflectance component enables applications like retexturing, material modification, object recolouring, palette extraction *etc.*, while shading component can be manipulated to perform shadow removal, glare removal, low light enhancement *etc.*

Illumination manipulation can be done at either global or local level. Global analysis involves techniques like histogram equalization, white balancing, gamma correction, light source direction computation, illumination spectra estimation *etc.* Local methods do neighbourhood analysis and involve methods performing spatially varying illumination and environmental map estimation techniques.

Illumination analysis has also been performed for applications like shadow, haze, underwater blur and glare removal. It is also carried out for image based rendering task like harmonization after compositing an object inside image and image relighting. In this paper we specifically focus on low light image enhancement methods and discuss them in detail in the next section.

Low Light Image Enhancement: One crucial category of illumination analysis is performed for the task of exposure correction in poorly lit images. Earlier methods in this field were based on either pixel intensity manipulation (statistically via histogram equalization [36, 40, 41], or individually via intensity curve manipulation [28, 60]) or tried to estimate illumination map based on Retinex theory [21, 35, 55]. Recently several deep learning based methods have been proposed which have been trained on paired image datasets with low light images and their enhanced counterparts. Some of these solutions have been proposed for raw inputs or require other camera parameters as input. We restrict our discussion to the works which assume standard *sRGB* image as input.

Gharbi et al. [22] proposed an image enhancement architecture by introducing deep bilateral filtering. Wei et al. [56] took inspiration from the retinex decomposition and proposed RetinexNet by first splitting the image into separate reflectance and illumination components. Later Zhang et al. [62] extended this idea by focusing on dark regions. [25] introduced a light weight zero-shot network which predicts global gamma correction parameters for the image which they apply iteratively to achieve enhancement. Afifi et al. [2] introduced a new dataset by changing exposure values from the information obtained from the raw images and train a sequential model on the basis of image Laplacian pyramid decomposition. Recently Jin et al. [32] proposed a multibranch model for separately handling the overexposure caused by the light sources in dark images. We use several such recent state-of-the-art methods for comparison in section 4.

3 QUATERNION FACTORIZED SIMULATED EXPOSURE FUSION

Figure 2 illustrates our QFSEF framework that can be divided into three sub-modules namely Layer Factorization, Stack Simulation and Exposure Fusion, which we individually discuss in detail below:
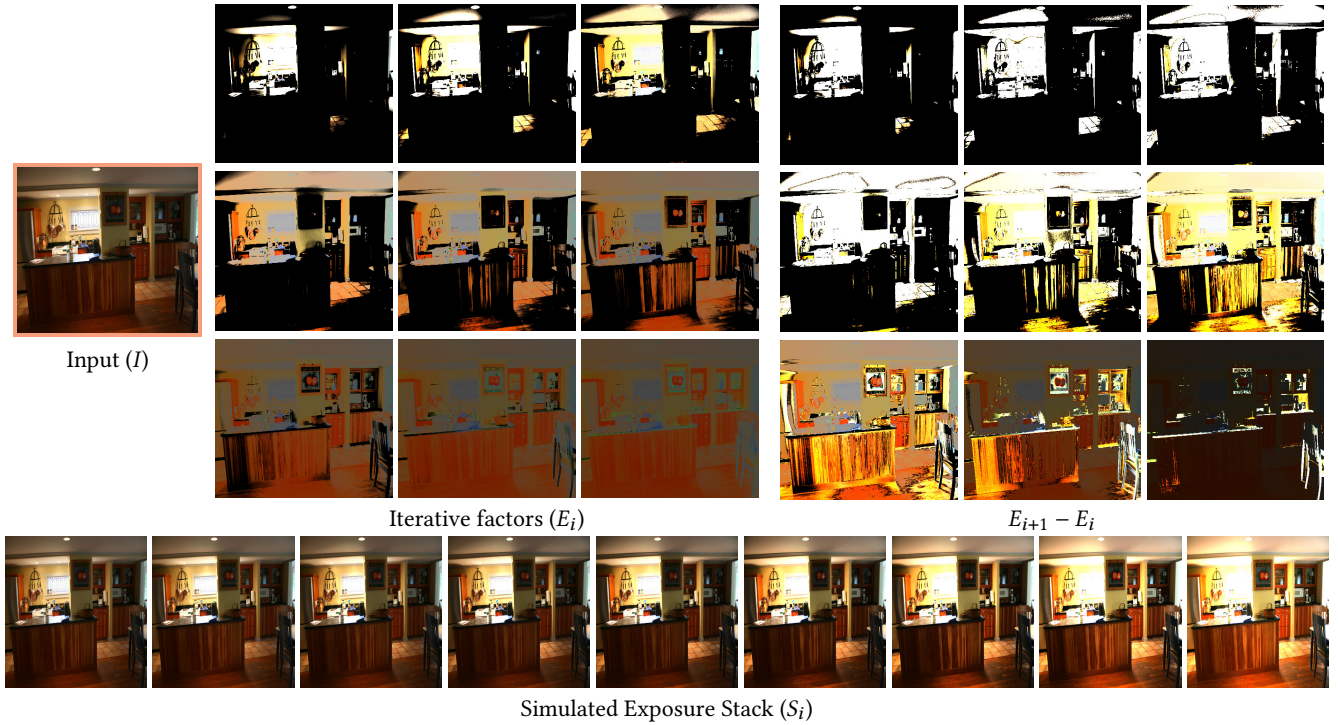
3.1 Layer Factorization

Our first sub-module performs factorization of the given input image into multiple layers. Factorization is carried out in such a way, so that the illumination characteristics are consistent in each layer. Specifically, we follow the *Dichromatic reflection model* [52] under which, for a given image I, the total irradiance at any pixel position x is the sum of diffuse and specular reflection components ($I_{diffuse}$ and $I_{specular}$):

$$I = R_d \cdot S_d + R_s \cdot S_s = I_{diffuse} + I_{specular}$$

where R_d and S_d (R_s and S_s) are diffuse (specular) material Reflectance and scene Shading respectively.

RPCA:. Specular component of the irradiance is sparse in nature and can be extracted from I via matrix factorization methods [1, 3, 26, 27, 61]. To this end we use Robust Principle Component

Figure 3: Layer Factorization: Input image (left) is iteratively split using quaternion RPCA into multiple factors (middle grid). Factor differences (right grid) highlights additional information captured in each successive factor. Last row shows the exposure stack simulated using these factors. All images have been luminance normalized (subsection 3.2) for visualization (more results available in supplementary).

Analysis (RPCA) [15] which has been used previously for various computer vision problems [10] like background separation, denoising, specularity removal, tracking *etc*. RPCA has a tractable convex approximation for the original NP-hard formulation:

$$I = \min_{A,E} A_{rank} + \lambda E_{sparse} \approx \min \|A\|_* + \lambda \|E\|_1 \quad s.t. \ I = A + E \quad (1)$$

Here E and A are estimations of the sparse specular and the remaining low-rank diffuse components respectively. Also $\|\|_*$ and $\|\|_1$ respectively stand for nuclear and L_1 norms which are used to approximate the original low-rank A_{rank} and sparse E_{sparse} matrices. λ is a positive constant for controlling the sparsity (specularity) of the result. Equation 1 can be solved using various optimization techniques [11]. Specifically, we employ the Augmented Lagrangian Method (ALM) [37] which uses Principal Component Pursuit (PCP) [15] for the optimization.

Quaternion Representation: We use the quaternion extension of PCP introduced by Chan and Yang [16] for instrument and voice signal separation in music clips. Quaternions are hypercomplex numbers which can be understood as an extension of the regular complex numbers in three dimensions with one scalar (r) and three vector components (i, j, k) *s.t.* $i^2 = j^2 = k^2 = -1$ and $i.j = k$; $j.k = i$; $k.i = j$; $i.k = -j$; $j.i = -k$ and $k.j = -1$. We first represent the normalized image I as a pure Quaternion matrix ($r=0$) by encoding each pixel as a *Versor* (*i.e.* a unit norm quaternion) with the color channels (RGB) encoded as the three quaternion vectors ($R \mapsto i$, $G \mapsto j$, $B \mapsto k$). The advantage of optimizing in quaternion space

compared to concatenated RGB matrices is that the quaternion PCP preserves both spectral and inter-channel phase information unlike Real (no phase) and Complex (spectral phase) variants [16, 20]. Thus the color channel information is better represented in this manner leading to improved factorization (empirically verified in section 4).

Iterative Factorization: In order to generate more than two simulated images we propose to apply the RPCA splitting iteratively on the factorized low-rank component by gradually reducing the sparsity constraint λ from $\lambda_{max} \rightarrow 1$. This can be understood as progressive subtraction of the specularity content from the factorized components till the total signal energy limits to zero. To achieve this, we gradually relax k in Equation 1.

$$I = E_1 + A_1 = E_1 + (E_2 + A_2) = E_1 + E_2 + E_3 + A_3 = ... = \sum_i^K E_i. \ (2)$$

Figure 3 shows an example of such factors and their successive differences (*i.e.* $E_{i+1} - E_i$) for visualization. Note how the initial factors contain predominantly scene highlights and light sources, followed by differently lit regions and finally shadows (loosely representing: sharp highlights $\rightarrow$ soft highlights $\rightarrow$ direct $\rightarrow$ indirect $\rightarrow ... \rightarrow$ soft shadows $\rightarrow$ dark shadows *etc*.). One can also look at these factors, especially the successive differences, as illumination consistent superpixel regions which are similar in their shading content instead of color. The scene hence gets factorized into multiple illumination consistent layers which can be globally processed with simple enhancement operations.

3.2 Stack Simulation

As the image layers obtained from the previous step posses similar optical characteristics, we can directly edit them with simple global image manipulation operators without introducing significant illumination artifacts. Based upon this idea, we simulate a virtual exposure stack from a single image for the purpose of low-light image enhancement task. We first factorize the input low-light image into K factors (E_i) by following the procedure in subsection 3.1. We post-process thus obtained layers by following three steps:

(i) **Layer grouping** *i.e.* based on layer signal energy merge with the next layer if lower than a set threshold ($\tau = 1\%$). This increases efficiency by simplifying the stack.

(ii) **Outliers removal** *i.e.* clip the pixels above and below certain percentiles (99.9 and 0.1) to the corresponding cutoff values. This controls extremely high and low valued noisy pixels.

(iii) **Luminance normalization** *i.e.* rescale the luminance of layers $\in [0, m_k]$ by normalizing the value in HSV color space. Here m_k is the signal energy in each factor which is estimated as the sum of all pixel values in that layer divided by sum of all m_ks *i.e.* $m_k = (\Sigma_x E_k)/(\Sigma_k \Sigma_x E_k)$. This enhances contrast while adhering to the original order of luminosity values of image regions.

After post-processing, we simulate our virtual exposure stack (S_i) by linearly combining the layers with I progressively:

$$S_{i+1} = (1 - \alpha)S_i + \alpha E_i, \quad where\ i \in [0, K]\ and\ S_0 = I. \quad (3)$$

Linear combination in Equation 3 helps in the gradual introduction of information from the successive layers into the low-lit image. This avoids sudden illumination jumps which might generate unnaturally lit images. Furthermore, it simulates the effect of slowly increasing the exposure time (or reducing shutter speed), leading to progressively brighter images with varying exposure values as shown in Figure 3.

3.3 Exposure Fusion

Fundamental task in exposure fusion is to assess the amount of information at each pixel in the stack and use it to merge all the images to render an enhanced image. It is a well understood problem with multiple solutions proposed in the current literature [11]. We build our framework by adapting multiple existing exposure fusion strategies which we discuss below:

Direct Fusion: As a simple baseline we process our last simulated image which contains the information from all the layers. Similar to subsection 3.2, we first percentile clip the outliers and then normalize the luminance. In order to remove the image noise arising from the sensor errors in the low-lit regions and rescaled in the dark layers, we use Block-Matching and 3D collaborative filtering (BM3D). BM3D is a non-local transform based denoising method which first spatially (2D) and then spectrally (1D) transforms similar image patches which are then denoised in this 3D domain via shrinkage [39]. After denoising, we rescale the image and use it as our directly fused enhanced result I_D.

Laplacian Pyramid Fusion: In their seminal work, Mertens et al. [38] introduced the concept of exposure fusion via a Laplacian pyramids from a bunch of Low Dynamic Range (LDR) to create a High Dynamic Range (HDR) output. They first estimate the image

quality at each pixel using three metrics: contrast $C_k(x)$, saturation $S_k(x)$ and well-exposedness $W_k(x)$ which respectively measure the spatial gradient magnitude, pixel color standard-deviation and Gaussian curve distance from the mid-tone value of 0.5. For each image a weight matrix is constructed by simply multiplying the three quality maps as $w_k(x) = C_k(x)^{\lambda_c}.S_k(x)^{\lambda_s}.W_k(x)^{\lambda_w}$ with λ_js tunable hyperparameters. The input images are then fused by merging their Laplacian pyramids weighed by the Gaussian pyramids of their respective normalized weights [12]. We ignore the contrast term as it is already handled by the Luminance normalization step during simulation in subsection 3.2 and empirically fix the hyperparameters values at $\lambda_c = 0$, $\lambda_s = 1$ and $\lambda_w = 2$. We denoise and rescale and denote the hence obtained enhanced result as I_L.

Generalized Random Walk Fusion: The third strategy that we experiment with is the Generalized Random Walk Fusion method by Shen et al. [50] who fuse multiple LDR images to construct a HDR image which is tone mapped to the enhanced result. They perform a global optimization using local contrast and color smoothness as two quality measures. They setup a Dirichlet problem [18] per stack image and solve them via global Generalized Random Walk algorithm [23]. They estimate dense probability maps representing significance of each pixel in the stack towards generating a combined enhanced image. We apply their algorithm on our simulated exposure stack and label our enhanced result as I_G, post denoising and rescaling.

4 EXPERIMENTS AND RESULTS

Implementation Details: We implement our framework using the *Matlab Quaternion Toolbox* [46]. Codes for various fusion strategies and BM3D denoising were adapted from their respective official releases [38, 39, 50]. We show the utility of our framework by comparing with the previous state-of-the-art single image low-light enhancement methods. We restrict ourselves to the solutions which directly work on *sRGB* images and do not require *RAW* inputs. Specifically we compare with RetinexNet by Wei et al. [56], DCENet by Guo et al. [25], LPNet by Afifi et al. [2] and UNIE by Jin et al. [32]. RetinexNet and LPNet are trained via direct supervision on LOL dataset [56] and simulated exposure stack from Adobe5k dataset [13] respectively. Whereas, DCENet and UNIE follow zero-shot and unpaired unsupervised learning adapted on SICE [14] and LOL [56] datasets respectively. We use the code and pretrained weights as shared by the authors' and reference the quantitative scores as reported in the literature.

Quantitative Results: We show both quantitative (PSNR-SSIM scores Q_i in Table 1) and qualitative results (Figure 6) on four datasets ($\mathcal{D}_i$) comprising of varying number of test images: LOLv1 testset (15) [56], LOLv2 testset (100) [58], SICE Part 2 testset (767) [14] as described by Guo et al. [25] and Adobe5k testset (5905 images) [2] . Adobe5k contains both over and under exposed images and provides 5 expert ground truth annotations. Afifi et al. [2] report both separate and combined results on these two categories for all the annotations. We follow the same procedure and show our overexposed, underexposed and combined scores averaged over all the expert annotations (for individual expert ground truth evaluations refer to the supplementary material).

| Methods | Testset → Trainset ↓ | LOLv1 test [56] (15) | LOLv2 test [58] (100) | SICE Part 2 [14] (767) | Adobe5k [2] (3543 + 2362 = 5905) | | | Average (S_a) | Generaliz -ability (S_g) |
					Overexp.	Underexp.	Complete		
RetinexNet[56]	LOLv1 train	16.77 - 0.46	15.47 - 0.56	15.99 - 0.53	11.06 - 0.60	12.49 - 0.62	11.63 - 0.61	12.19 - 0.60	12.18 - 0.60
DCENet [25]	SICE Part 1	14.86 - 0.59	<u>20.54</u> - <u>0.78</u>	16.57 - 0.59	11.02 - 0.52	14.96 - 0.59	12.60 - 0.55	13.17 - 0.56	12.74 - 0.55
LPNet [2]	raw Adobe5k	15.3 - 0.56	16.38 - 0.53	14.55 - 0.50	**19.35 - 0.74**	<u>19.69</u> - **0.74**	**19.48 - 0.74**	**18.87 - 0.71**	14.77 - 0.51
UNIE [32]	LOLv1 train	**21.52** - <u>0.76</u>	**25.53 - 0.88**	13.72 - 0.46	16.93 - 0.66	15.65 - 0.60	16.41 - 0.64	16.26 - 0.62	16.42 - 0.62
I_D (Direct)	_	<u>20.39</u> - **0.77**	19.12 - 0.67	<u>16.82</u> - **0.62**	17.90 - <u>0.71</u>	**19.87** - <u>0.72</u>	18.69 - <u>0.71</u>	18.49 - <u>0.70</u>	<u>18.49</u> - **0.70**
I_L (Laplacian)	_	19.28 - 0.75	18.16 - 0.67	**17.75** - <u>0.60</u>	15.60 - 0.65	17.94 - 0.69	16.78 - 0.67	16.92 - 0.66	16.92 - 0.66
I_G (GRWF)	_	17.72 - 0.70	19.01 - 0.69	15.64 - 0.56	<u>18.71</u> - <u>0.71</u>	19.34 - 0.70	<u>18.96</u> - <u>0.71</u>	<u>18.58</u> - 0.69	**18.58** - <u>0.69</u>

Table 1: Quantitative comparison: We evaluate our simulated exposure stack generation scheme over 5 datasets using results from our 3 exposure fusion strategies (I_D, I_L, I_G). Each tuple represents PSNR-SSIM scores (higher is better). S_a is average score weighted by testset size and S_g is method's *Generalizability* score computed as weighted average leaving out the testset corresponding to its supervision dataset. Best score is in **boldfaced** and second best is <u>underlined</u>.
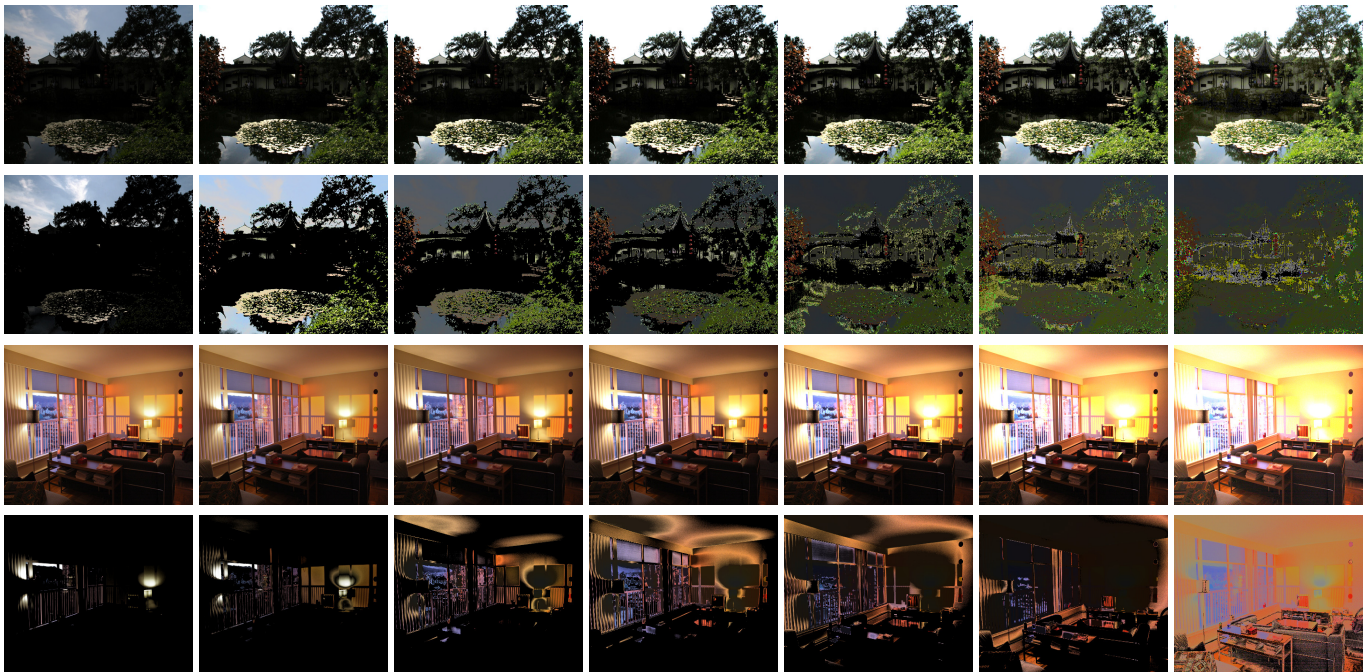

Figure 4: Stack Simulation: This figure shows Simulated Exposure Stack and the underlying quaternion RPCA decomposed factors for two types of scenes: outdoor naturally lit scene with single light source (top) and indoors artificially lit room with multiple illuminants (bottom).

For overall performance, we report average scores (S_a) weighted by testset cardinality *i.e.* $\Sigma_i(|\mathcal{D}_i|Q_i)/\Sigma_i|\mathcal{D}_i|$ thus assigning equal importance to all test images across datasets. Furthermore in order to gauge the cross-dataset generalizability of learning based methods, we also calculate the average scores by leaving out the test images of the dataset on which the respective model was trained (*e.g.* leaving out LOLv1 testset for RetinexNet and computing weighted mean over the rest). We report this metric as the *Generalizability Index* (S_g) of the method and list them in the last column. As our method is a non-data driven optimization algorithm, last two columns are same in our case *i.e.* $S_a = S_g$. We provide evaluation scores for all three of our fusion strategies in the three last rows (I_D, I_L and I_G).

As can be observed from Table 1, our results are frequently ranked best or second best on multiple datasets even without having learned any data-driven prior via training. As expected, all four learning based methods perform well on the datasets which are similar to their respective training sets but suffer significant performance degradation when dataset domain shifts. The effect pronounced in the case of smaller models like RetinexNet and DCENet. (Note that datasets LOLv1 and LOLv2 are similar than the rest). In overall performance S_a our method ranks second just behind LPNet [2]. Note that LPNet was trained on simulated images from the raw data in Adobe5k. Hence their performance on the corresponding testset is also higher which contributes a high weight to the overall score S_a. In order to ameliorate this, we use
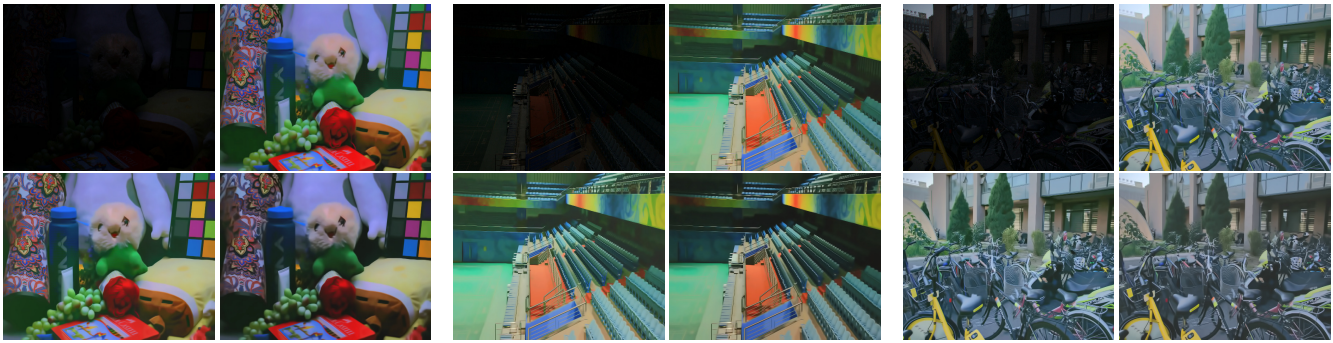
Figure 5: Results: Our low light image enhancement results on three scenes using our three exposure fusion strategies (subsection 3.3). Clockwise from top left: I, I_D, I_L and I_G respectively. Test images taken from LOL dataset [56, 58].

| Input | DCENet [25] | LPNet [2] | UNIE [32] | Our results (I_L) | Ground Truth |
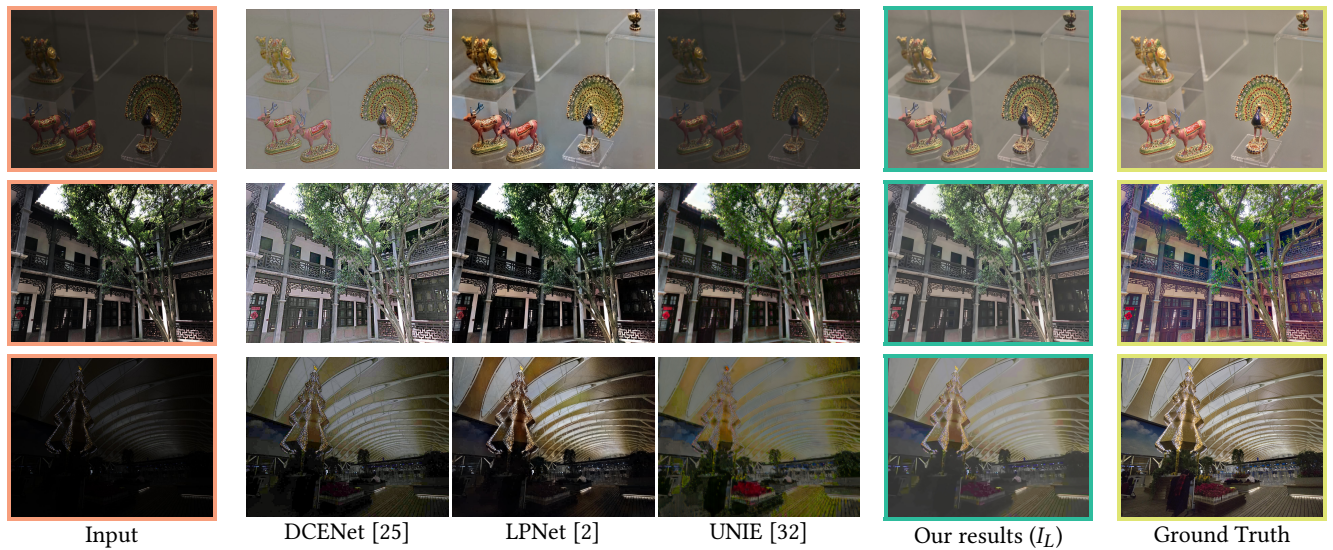
Figure 6: Qualitative Comparison: Low light enhancement results comparison between DCENet [25], LPNet [2], UNIE [32] and our method I_L. DCENet [25] leads to desaturated colors while LPNet [2] and UNIE [32] fail to properly illuminate some regions. Our method achieves good enhancement without significant color degradation.

the Generalizability Index score S_g by leaving out significance of respective training dataset test images. Under this metric all of our strategies perform better than the current state-of-the-art solutions indicating our wider generalizability and utility.

Qualitative Results: The first step of our method *i.e.* Quaternion RPCA factorization, is not limited to dimly lit scenes and can be applied on a variety of images as shown in Figure 4. We show our resultant simulated exposure stack from a single under-, well- and over- exposed image input. in Figure 5 we show three sample enhancement results for a low light image using all three of our strategies (I_D, I_L and I_G). Out of these three, I_D is the fastest but does not integrate contrast from the initial layers quite well. On the other hand, I_G is the slowest but has better contrast. I_L though has relatively lesser quantitative score but strikes a balance between both the scenarios. any of these strategies can be used for the task based on user preferences. We also gauge the perceptual quality of our results (I_L) by comparing against previous SOTA in Figure 6.

It can be observed that DCENet [25] brightens the image but dulls the color information and the resultant images are grainy. LPNet [2] leaves some dark shadows and over saturates colors in some regions. UNIE [32] results have good color but are blurred and dark. They are also inconsistent depending on the scene type. Our method achieves balanced enhanced results with good contrast and are reported in the last column.

Ablation Analysis: We validate our design choices by performing two different ablation analyses on LOLv1 testset images. First we compare four different variants of our system. Variant v_1 is performed by swapping quaternion RPCA with real RPCA on the color channels concatenated representation of the image. Variant v_2 is designed to signify the impact of denoising and is formed by skipping the CMB3D denoising step in the original pipeline. v_3 is very similar to v_4 only with luminance normalization step replaced with normal range normalization whereas v_4 represents our complete system. As seen in Table 2 and Figure 7, v_1 fails to capture the color
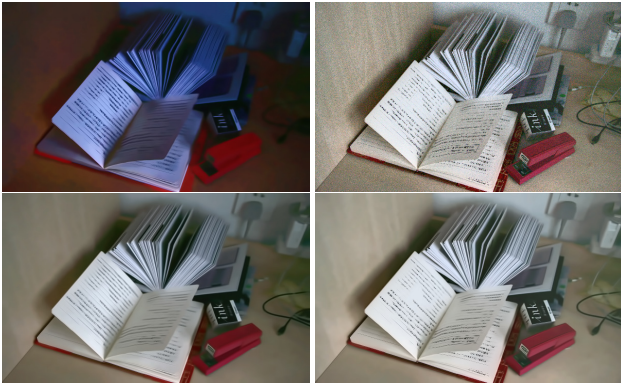
Figure 7: Ablation: Illustrative low light enhancement results for our four variants v_1, v_2 (top) and v_3, v_4 (bottom) in two rows respectively. v_1 optimizes in real space and hence does poor color preservation. Without denoising, v_2 contains salt-pepper noise. Note slight over smoothing of the text in the notebook and dark corners in v_3 results compared to our final design choices in v_4.
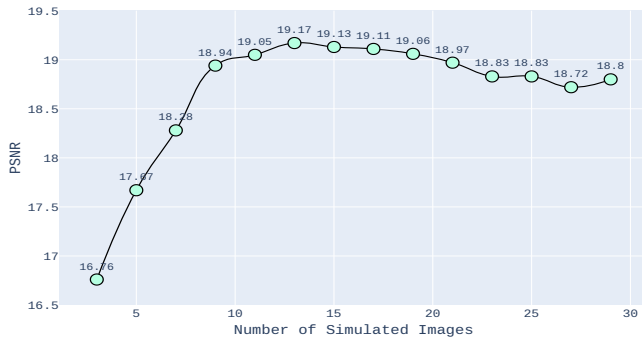
Figure 8: Ablation on k parameter: The graph above shows the effect of different number of simulated images on the mean PSNR of the three strategies. After swift increase initially, performance plateaus after $k = 10$ with a slight degradation towards the end. We choose the middle plateau value of $k = 15$ for our experiments.

information properly, whereas v_2 contains amplified dark channel noise prevalent in the low-light images. v_3 performs quite well quantitatively but does over smoothing in high frequency parts in certain scenes. Overall v_4 presents a balanced approach with good numerical and perceptual accuracy.

Our second ablation experiment is done to analyze the effect of number of simulated images K. We gradually increase the value of K (3, 5, 7, ...29) and observe the mean PSNR scores computed on our ablation set. As can be observed from Figure 8, the scores after increasing drastically for first few values, plateaus in the range $K = [10\text{-}15]$. This happens because the fewer number of factors do not represent illumination consist regions in the image properly and hence fail to properly enhance the poorly lit pixel values. On the other hand higher values of K lead to several dimly lit factors which contain sensor noise, thus degrading the quality of stack input in the exposure fusion step and thereby reducing the quality of generated results. To enable better user control over the scene, we fix $K = 15$ in all our experiments.

Variants $\longrightarrow$	real RPCA v_1	w/o Denoise v_2	w/o LNorm. v_3	Full v_4
I_D	12.83 - 0.5	18.35 - 0.6	20.86 - 0.75	20.39 - 0.77
I_L	14.54 - 0.56	18.14 - 0.59	20.11 - 0.76	19.28 - 0.75
I_G	12.66 - 0.48	15.83 - 0.53	17.48 - 0.69	17.72 - 0.70

Table 2: Ablation using system variants: We show PSNR-SSIM ($\uparrow$) scores on 15 LOLv1 [56] testset images for our four system variants *i.e.* using real RPCA (v_1), without denoising (v_2), without luminance normalization (v_3) and our complete version (v_4). The performance gradually improves for each step empirically validating our design choices. Note that although we achieve higher scores with v_3 but it leads to oversmoothing of edges especially in high frequency regions (see Figure 7).

5 CONCLUSION AND FUTURE DIRECTIONS

To summarize, in this work we have presented a novel exposure stack simulation method and applied it to the task of low-light image enhancement. We extend the Retinex theory by proposing a Robust Retinex Decomposition formulation. To this end we present a new way of image factorization by first representing the image pixels in the quaternion space as a pure quaternion vector and then apply Robust Principal Component Analysis to obtain sparse factors representing the specular information in the image. We propose a scheme to simulate a virtual exposure stack from a single image by iterative factorization and adapt existing exposure fusion strategies to generate an enhanced well-lit image. Our results show good qualitative and quantitative performance on multiple datasets, especially exhibiting marked generalization performance even when compared with the contemporary state-of-the-art deep learning based solutions for this task.

In the end, we would like to point out the our proposed factorization technique can assist several related image based rendering problems beyond the low light image enhancement task like selective relighting, shadow removal, white balancing, object compositing, image harmonization *etc*. Furthermore, we can extend our method by assembling an end-to-end deep model for the three sub-modules *i.e.* factorization, simulation and fusion. In future we would like to pursue these directions.

REFERENCES

[1] Amir Adler, Michael Elad, Yacov Hel-Or, and Ehud Rivlin. 2013. Sparse coding with anomaly detection. In *International Workshop on Machine Learning for Signal Processing (MLSP)*.

[2] Mahmoud Afifi, Konstantinos G. Derpanis, Björn Ommer, and M. S. Brown. 2021. Learning Multi-Scale Photo Exposure Correction. *Computer Vision and Pattern Recognition (CVPR)* (2021).

[3] Yasushi Akashi and Takayuki Okatani. 2016. Separation of reflection components by sparse non-negative matrix factorization. *Computer Vision and Image Understanding* 146, C (2016).

[4] Yagiz Aksoy, Tunç Ozan Aydin, Aljoscha Smolic, and Marc Pollefeys. 2017. Unmixing-Based Soft Color Segmentation for Image Manipulation. *ACM Transactions on Graphics* 36 (2017).

[5] Yagiz Aksoy, Tae-Hyun Oh, Sylvain Paris, Marc Pollefeys, and Wojciech Matusik. 2018. Semantic soft segmentation. *ACM Transactions on Graphics (TOG)* 37 (2018).

[6] Dimitrios S. Alexiadis and George D. Sergiadis. 2009. Estimation of Motions in Color Image Sequences Using Hypercomplex Fourier Transforms. *IEEE Transactions on Image Processing* 18 (2009).

[7] Dawit Assefa, Lalu Mansinha, Kristy F. Tiampo, Henning Rasmussen, and Kenzu Abdella. 2010. Local quaternion Fourier transform and color image texture analysis. *Signal Processing* 90, 6 (2010).

[8] Harry G. Barrow and Jay M. Tenenbaum. 1978. Recovering Intrinsic Scene Characteristics from Images. In *A. Hanson and E. Riseman, Eds., Computer Vision Systems.*

[9] Nicolas Le Bihan and Stephen J. Sangwine. 2003. Quaternion principal component analysis of color images. *International Conference on Image Processing (ICIP)* 1 (2003).

[10] Thierry Bouwmans, Sajid Javed, Hongyan Zhang, Zhouchen Lin, and Ricardo Otazo. 2018. On the Applications of Robust PCA in Image and Video Processing. *Proc. IEEE* 106 (2018).

[11] Thierry Bouwmans, Andrews Sobral, Sajid Javed, Soon Ki Jung, and El-Hadi Zahzah. 2017. Decomposition into low-rank plus additive matrices for background/foreground separation: A review for a comparative evaluation with a large-scale dataset. *Computer Science Review* 23 (2017), 1–71.

[12] Peter J. Burt. 1985. Pyramid-based Computer Graphics the Pyramid Representation and Computer Graphics.

[13] Vladimir Bychkovsky, Sylvain Paris, Eric Chan, and Frédo Durand. 2011. Learning Photographic Global Tonal Adjustment with a Database of Input / Output Image Pairs. In *Computer Vision and Pattern Recognition (CVPR).*

[14] Jianrui Cai, Shuhang Gu, and Lei Zhang. 2018. Learning a Deep Single Image Contrast Enhancer from Multi-Exposure Images. *IEEE Transactions on Image Processing* 27, 4 (2018).

[15] Emmanuel J. Candès, Xiaodong Li, Yi Ma, and John Wright. 2011. Robust Principal Component Analysis? *J. ACM* 58, 3, Article 11 (2011).

[16] Tak-Shing T. Chan and Yi-Hsuan Yang. 2016. Complex and Quaternionic Principal Component Pursuit and Its Application to Audio Separation. *IEEE Signal Processing Letters* 23 (2016).

[17] Yi Chang, Luxin Yan, Tao Wu, and Sheng Zhong. 2016. Remote Sensing Image Stripe Noise Removal: From Image Decomposition Perspective. *IEEE Transactions on Geoscience and Remote Sensing* 54 (2016).

[18] Peter G. Doyle and J. Laurie Snell. 1984. *Random Walks and Electric Networks.* Number Book 22 in Carus Mathematical Monographs. Mathematical Association of America.

[19] Todd A. Ell and Stephen J. Sangwine. 2007. Hypercomplex Fourier Transforms of Color Images. *IEEE Transactions on Image Processing* 16 (2007).

[20] Todd A. Ell and Stephen J. Sangwine. 2007. Hypercomplex Fourier Transforms of Color Images. *IEEE Transactions on Image Processing* 16 (2007).

[21] Xueyang Fu, Delu Zeng, Yue Huang, Yinghao Liao, Xinghao Ding, and John Paisley. 2016. A fusion-based enhancing method for weakly illuminated images. *Signal Processing* 129 (2016).

[22] Michaël Gharbi, Jiawen Chen, Jonathan T. Barron, Samuel W. Hasinoff, and Frédo Durand. 2017. Deep bilateral learning for real-time image enhancement. *ACM Transactions on Graphics (TOG)* 36 (2017).

[23] L. Grady. 2006. Random Walks for Image Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 28, 11 (2006).

[24] Shuhang Gu, Lei Zhang, Wangmeng Zuo, and Xiangchu Feng. 2014. Weighted Nuclear Norm Minimization with Application to Image Denoising. *Computer Vision and Pattern Recognition (CVPR)* (2014).

[25] Chunle Guo, Chongyi Li, Jichang Guo, Chen Change Loy, Junhui Hou, Sam Kwong, and Cong Runmin. 2020. Zero-reference deep curve estimation for low-light image enhancement. *Computer Vision and Pattern Recognition (CVPR)* (2020).

[26] Jie Guo, Zuojian Zhou, and Limin Wang. 2018. Single Image Highlight Removal with a Sparse and Low-Rank Reflection Model. In *European Conference on Computer Vision (ECCV).*

[27] Xiaojie Guo, Xiaochun Cao, and Yi Ma. 2014. Robust Separation of Reflection from Multiple Images. In *Computer Vision and Pattern Recognition (CVPR).*

[28] Charles Hessel. 2019. Simulated Exposure Fusion. *Image Processing On Line* 9 (2019).

[29] Xiaodi Hou, Jonathan Harel, and Christof Koch. 2012. Image Signature: Highlighting Sparse Salient Regions. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 34 (2012).

[30] David E Jacobs, Jongmin Baek, and Marc Levoy. 2012. Focal stack compositing for depth of field control. *Stanford Computer Graphics Laboratory Technical Report* 1, 1 (2012).

[31] Zhigang Jia, Michael K. Ng, and Guang-Jing Song. 2019. Robust quaternion matrix completion with applications to image inpainting. *Numerical Linear Algebra with Applications* 26 (2019).

[32] Yeying Jin, Wenhan Yang, and Robby T Tan. 2022. Unsupervised Night Image Enhancement: When Layer Decomposition Meets Light-Effects Suppression. *European Conference on Computer Vision (ECCV)* (2022).

[33] Aliaa T. Kamal and Moumen T. El-Melegy. 2017. Color Image Processing Using Reduced Biquaternions with Application to Face Recognition in a PCA Framework. *International Conference on Computer Vision Workshops (ICCVW)* (2017).

[34] Ivan Kyrchei. 2013. Explicit representation formulas for the minimum norm least squares solutions of some quaternion matrix equations. *Linear Algebra Appl.* 438 (2013).

[35] Edwin Herbert Land. 1977. The retinex theory of color vision. *Scientific American* 237 6 (1977).

[36] Chulwoo Lee, Chul Lee, and Chang-Su Kim. 2013. Contrast Enhancement Based on Layered Difference Representation of 2D Histograms. *IEEE Transactions on Image Processing* 22, 12 (2013).

[37] Zhouchen Lin, Minming Chen, and Yi Ma. 2010. The Augmented Lagrange Multiplier Method for Exact Recovery of Corrupted Low-Rank Matrices. *Mathematical Programming* 9 (09 2010).

[38] Tom Mertens, Jan Kautz, and Frank Van Reeth. 2009. Exposure Fusion: A Simple and Practical Alternative to High Dynamic Range Photography. *Computer Graphics Forum* 28 (2009).

[39] Ymir Mäkinen, Lucio Azzari, and Alessandro Foi. 2020. Collaborative Filtering of Correlated Noise: Exact Transform-Domain Variance for Improved Shrinkage and Patch Matching. *IEEE Transactions on Image Processing* 29 (2020).

[40] S.M. Pizer, R.E. Johnston, J.P. Ericksen, B.C. Yankaskas, and K.E. Muller. 1990. Contrast-limited adaptive histogram equalization: speed and effectiveness. In *Visualization in Biomedical Computing.*

[41] Stephen M. Pizer, Elton Philip Amburn, John D. Austin, Robert Cromartie, Ari Geselowitz, Trey Greer, Bart M. ter Haar Romeny, and John B. Zimmerman. 1987. Adaptive histogram equalization and its variations. *Graphical Models graphical Models and Image Processing computer Vision, Graphics, and Image Processing* 39 (1987).

[42] Saurabh Saini and P. J. Narayanan. 2018. Semantic Priors for Intrinsic Image Decomposition. In *British Machine Vision Conference 2018, (BMVC).*

[43] Saurabh Saini and P. J. Narayanan. 2019. Semantic Hierarchical Priors for Intrinsic Image Decomposition. *ArXiv* abs/1902.03830 (2019).

[44] Saurabh Saini, Parikshit Sakurikar, and P. J. Narayanan. 2016. Intrinsic image decomposition using focal stacks. In *Indian Conference on Computer Vision, Graphics and Image Processing (ICVGIP).*

[45] S.J. Sangwine. 2000. Colour image filters based on hypercomplex convolution. *Vision, Image and Signal Processing* 147 (2000). Issue 2.

[46] Steve Sangwine and Nicolas Le Bihan. 2022. *Quaternion toolbox for Matlab. Version 3.* http://qtfm.sourceforge.net/

[47] Stephen J. Sangwine. 1998. Colour image edge detector based on quaternion convolution. *Electronics Letters* 34 (1998).

[48] Stephen J. Sangwine and Nicolas Le Bihan. 2006. Quaternion singular value decomposition based on bidiagonalization to a real or complex matrix using quaternion Householder transformations. *Appl. Math. Comput.* 182, 1 (2006).

[49] Boris Schauerte and Rainer Stiefelhagen. 2012. Quaternion-Based Spectral Saliency Detection for Eye Fixation Prediction. In *European Conference on Computer Vision (ECCV).*

[50] Rui Shen, Irene Cheng, Jianbo Shi, and Anup Basu. 2011. Generalized Random Walks for Fusion of Multi-Exposure Images. *IEEE Transactions on Image Processing* 20 (2011).

[51] Özlem Nurcan Subakan and Baba C. Vemuri. 2010. A Quaternion Framework for Color Image Smoothing and Segmentation. *International Journal of Computer Vision* 91 (2010).

[52] Shoji Tominaga. 1994. Dichromatic reflection models for a variety of materials. *Color Research and Application* 19 (1994).

[53] Richard Tucker and Noah Snavely. 2020. Single-View View Synthesis With Multiplane Images. *Computer Vision and Pattern Recognition (CVPR)* (2020).

[54] Georgios Tzimiropoulos, Stefanos Zafeiriou, and Maja Pantic. 2012. Subspace Learning from Image Gradient Orientations. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 34 (2012).

[55] Shuhang Wang, Jin Zheng, Hai-Miao Hu, and Bo Li. 2013. Naturalness Preserved Enhancement Algorithm for Non-Uniform Illumination Images. *IEEE Transactions on Image Processing* 22 (2013).

[56] Chen Wei, Wenjing Wang, Wenhan Yang, and Jiaying Liu. 2018. Deep Retinex Decomposition for Low-Light Enhancement. In *British Machine Vision Conference (BMVC).*

[57] Dongpo Xu, Yili Xia, and Danilo P. Mandic. 2016. Optimization in Quaternion Dynamic Systems: Gradient, Hessian, and Learning Algorithms. *IEEE Transactions on Neural Networks and Learning Systems* 27 (2016).

[58] Wenhan Yang, Wenjing Wang, Haofeng Huang, Shiqi Wang, and Jiaying Liu. 2021. Sparse Gradient Regularized Deep Retinex Network for Robust Low-Light Image Enhancement. *IEEE Transactions on Image Processing* 30 (2021).

[59] Qilin Yin, Jinwei Wang, Xiangyang Luo, Jiangtao Zhai, Sunil Kr. Jha, and Yun Qing Shi. 2019. Quaternion Convolutional Neural Network for Color Image Classification and Forensics. *IEEE Access* 7 (2019).

[60] Lu Yuan and Jian Sun. 2012. Automatic Exposure Correction of Consumer Photographs. In *European Conference on Computer Vision (ECCV).*

[61] Wuming Zhang, Xi Zhao, Jean-Marie Morvan, and Liming Chen. 2019. Improving Shadow Suppression for Illumination Robust Face Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 41, 3 (2019).

[62] Yonghua Zhang, Jiawan Zhang, and Xiaojie Guo. 2019. Kindling the Darkness: A Practical Low-light Image Enhancer. *ACM International Conference on Multimedia* (2019).

[63] Xiaowei Zhou, Can Yang, Hongyu Zhao, and Weichuan Yu. 2014. Low-Rank Modeling and Its Applications in Image Analysis. *ACM Computing Surveys (CSUR)* 47 (2014).