The Half-quadratic Splitting (HQS) Method for Shadow Removal

Yujie Gao

Abstract—Shadows caused by uneven illumination can significantly degrade image quality and hinder computer vision tasks such as image segmentation and object recognition. While current approaches, including pixel-level computations and learning-based methods, have shown satisfactory results and high automation, they often involve complex algorithms or require extensive training data. In this work, we propose a simple yet effective framework for shadow removal by formulating it as an inverse problem. Our approach considers shadows as Gaussian noise in the Y (luminance) channel and restores the shadow-free image using the Half-Quadratic Splitting (HQS) method. Experimental results demonstrate the feasibility and effectiveness of our algorithm in removing shadows and improving image quality, as evidenced by quantitative metrics, PSNR, SSIM and RMSE.

Index Terms—Computational Imaging, Shadow Removal, HQS

1 Introduction

Capturing high-quality images under varying lighting conditions is a common challenge in computational imaging. Shadows, caused by uneven illumination, can degrade image quality by reducing visibility and introducing unwanted artifacts. These artifacts, as well as changes due to local variation in the intensity or color of the illumination also make it more difficult to implement image segmentation, object recognition and other basic tasks for computer vision.

The task of shadow removal is to find the best estimate of the true image without shadow, x, based on an image with shadows, b. In some case, the shadow mask is given, while in most cases, we aren't able to obtain the accurate mask. In addition, we lack information of the illumination, which greatly affects the formation of shadows.

For efficient and convenient shadow removal, we think of it as an inverse problem and develop a simple framework using the Half-quadratic Splitting (HQS) method to find an acceptable solution. Performance of the algorithm is evaluated using three metrics, RMSE, PSNR and SSIM. Further suggestions on how the method could be improved are also discussed.

2 Related Work

Several existing methods address shadow removal and contrast enhancement. Notable approaches include histogram equalization [1], Retinex-based algorithms [2], [3], and deep learning techniques [4].

He-related methods involve applying to each pixel the histogram equalization mapping based on the pixels in a region surrounding that pixel, which is invented by Ketcham et al. [5], [6], [7]. AHE (Adaptive Histogram Equalization) [1] achieves improvement in speed and quality and operates on local regions, making it effective in enhancing details and mitigating the effects of varying illumination conditions across an image.

The Retinex theory posits that the visual system separates the influence of illumination and reflectance when perceiving an image. The illumination change can be modelled as an independent scaling of sensor responses in each channel (a so-called diagonal model of illumination change and consider a single surface which is partially in shadow. A retinex path entirely in shadow or entirely out of shadow gives the same results at each pixel since it is merely calculating ratios within a channel [3].

Many works have been put in shadow detection. Khan et al. [4] propose a framework which automatically learns the most relevant features in a supervised manner using multiple convolutional deep neural networks (ConvNets). Using the detected shadow masks, they propose a Bayesian formulation to accurately extract shadow matte and subsequently remove shadows. Deep learning methods have more adaptability to varied tasks compared with previous pixel-level computation.

Based on previous works and understandings, we seek to propose a simple and straightforward modeling of shadow formation and solve the problem using the HQS method.

3 Proposed Method

In this section, we firstly review the Image Formation Model for deconvolution, and then propose the inspired Shadow Formation Model. To solve the inverse problem, we take a Bayesian perspective of the problem and implement regularized shadow removal with Half-quadratic Splitting (HQS).

3.1 Image Formation for Deconvolution

For deconvolution, we are given a known shift-invariant point spread function (PSF and an optical low-pass filter in this case) and a 2D measured image. The goal for deconvolution is to find an estimate of the latent image from

• *Y. Gao is with the Institute for Computational and Mathematical Engineering (ICME), Stanford University, Stanford, CA, 94305. E-mail: gaoyj99@stanford.edu*

the blurry, possibly with noise, measurement. Having a 2D image x and a shift-invariant PSF c, the measurement b can be formed as

$$b = c * x + \eta \tag{1}$$

Here, η is an additive, signal-independent noise term.

Using the convolution system, Equation 1 can be written as the following

$$b = \mathcal{F}^{-1}\{\mathcal{F}\{c\} \cdot \mathcal{F}\{x\}\} + \eta \tag{2}$$

where $\cdot$ is the element-wise product.

To describe the 2D image using algebraic notation, we use the following expression

$$b = c * x \leftrightarrow \mathbf{b} = \mathbf{C}\mathbf{x} \tag{3}$$

where $\mathbf{x}, \mathbf{b} \in \mathbb{R}^n$ are the vectorized forms of the images x and b, and $\mathbf{C} \in \mathbb{R}^{n \times n}$ is a square Toeplitz matrix implementing a shift invariant convolution with c [8].

3.2 Image Formation for Shadow Removal

Shadows are formed in the parts of a scene where obstructing objects prevent direct light from a source from reaching those areas. The shadow regions, however, are illuminated by ambient light. Typically, shadows can be divided into two major classes (self-shadows and cast shadows). The cast shadow can be further divided into umbra and penumbra. and based on intensity, shadows can be divided into hard and soft shadows [9].

Based on the formation theories of shadows, we propose a Shadow Formation Model and view shadows as noise in the luminance channel, which means that main removal process will be implemented in the Y (luma) channel in YCbCr color space.

In analogy with the image formation for deconvolution, the image formation for shadow removal can be expressed as

$$\mathbf{b} = \mathbf{x} + \boldsymbol{\eta} \tag{4}$$

$$\mathbf{x}_i \sim \mathcal{N}(\mathbf{x}_i, 0),\ \boldsymbol{\eta}_i \sim \mathcal{N}(\mu, \sigma^2)(\mu < 0) \tag{5}$$

$$\mathbf{b}_i \sim \mathcal{N}(\mathbf{x}_i + \mu, \sigma^2) \tag{6}$$

where $\mathbf{x}, \mathbf{b} \in \mathbb{R}^n$ are the vectorized forms of the desired shadow-free image x and the measurement (observation/image with shadow) b. η represents shadow or Gaussian noise.

The probability of observation i and the joint probability of all observations will be

$$p(\mathbf{b}_i|\mathbf{x}_i, \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(\mathbf{b}_i - \mathbf{x}_i - \mu)^2}{2\sigma^2}} \tag{7}$$

$$p(\mathbf{b}|\mathbf{x}, \mu, \sigma) = \prod_{i=1}^{n} p(\mathbf{b}_i|\mathbf{x}_i, \mu, \sigma) \propto e^{-\frac{\|\mathbf{b} - \mathbf{x} - \boldsymbol{\mu}\|^2}{2\sigma^2}} \tag{8}$$

Since we want to get the posterior probability, we can have the following probability using Bayes' rule and get the Maximum-a-posterior (MAP) solution.

$$p(\mathbf{x}|\mathbf{b}, \mu, \sigma) = \frac{p(\mathbf{b}|\mathbf{x}, \mu, \sigma)p(\mathbf{x})}{p(\mathbf{b})} \propto p(\mathbf{b}|\mathbf{x}, \mu, \sigma)p(\mathbf{x}) \tag{9}$$

$$\mathbf{x}_{MAP} = \arg\min_{\mathbf{x}} \frac{1}{2\sigma^2}\|\mathbf{b} - \mathbf{x} - \boldsymbol{\mu}\|^2 + \Psi(\mathbf{x}) \tag{10}$$

3.3 Regularized Shadow Removal with HQS

To solve the inverse problem, we use the Half-quadratic Splitting (HQS) Method with TV prior. Based on derivation, we can reformulate the objective and remove the constraints using the penalty term.

$$\text{minimize}_{\mathbf{x}} \frac{1}{2}\|\mathbf{b} - \mathbf{x} - \boldsymbol{\mu}\|^2 + \lambda\Psi(\mathbf{D}\mathbf{x}) \tag{11}$$

$$L_\rho(\mathbf{x}, \mathbf{z}) = f(\mathbf{x}) + g(\mathbf{z}) + \frac{\rho}{2}\|\mathbf{D}\mathbf{x} - \mathbf{z}\|^2 \tag{12}$$

$$f(\mathbf{x}) = \frac{1}{2}\|\mathbf{b} - \mathbf{x} - \boldsymbol{\mu}\|^2,\ g(\mathbf{z}) = \lambda\Psi(\mathbf{D}\mathbf{x}) \tag{13}$$

Follow the alternating gradient descent approach and iteratively update $\mathbf{x}$ and $\mathbf{z}$ using the following rules

$$\mathbf{x} \leftarrow \text{prox}_{\|\cdot\|_2,\rho}(\mathbf{z}) = \arg\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{b} - \mathbf{x} - \boldsymbol{\mu}\|^2 + \frac{\rho}{2}\|\mathbf{D}\mathbf{x} - \mathbf{z}\|^2 \tag{14}$$

$$\mathbf{x} \leftarrow (\rho\mathbf{D}^T\mathbf{D} + \mathbf{I})^{-1}(-\boldsymbol{\mu} + \mathbf{b} + \rho\mathbf{D}^T\mathbf{z}) \tag{15}$$

$$\mathbf{z} \leftarrow \text{prox}_{\|\cdot\|_1,\rho}(\mathbf{D}\mathbf{x}) = \arg\min_{\mathbf{z}} \lambda\|\mathbf{z}\|_1 + \frac{\rho}{2}\|\mathbf{D}\mathbf{x} - \mathbf{z}\|^2 \tag{16}$$

For $\mathbf{z}$, we use element-wise soft thresholding operator $\mathcal{S}_\kappa(\cdot)$ for efficient update.

3.4 Implementation

3.4.1 *Estimation of $\boldsymbol{\mu}$*

Different from deconvolution, the mean of noise for shadow should be negative, based on the formation of shadow. The input to the algorithm is the shadow image and the shadow mask, which means that μ is not given. We can simply estimate μ from the shadow area and non-shadow area in the given shadow image.

$$\mu = \mu_{sh} - \mu_{nonsh} \tag{17}$$

Then, we use the computed μ in the iterative process. We mention that we view the formation of shadow as additive Gaussian noise with negative mean. To support this idea, we use the shadow image and shadow-free image in the data triplets to compute the distribution of the difference (in Y channel) in the shadow area (Fig. 2). As can be seen in the figure, the distribution of difference (treated as noise) looks like Gaussian distribution and its mean is negative as expected.

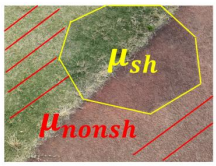

Fig. 1. Example image for μ computation.

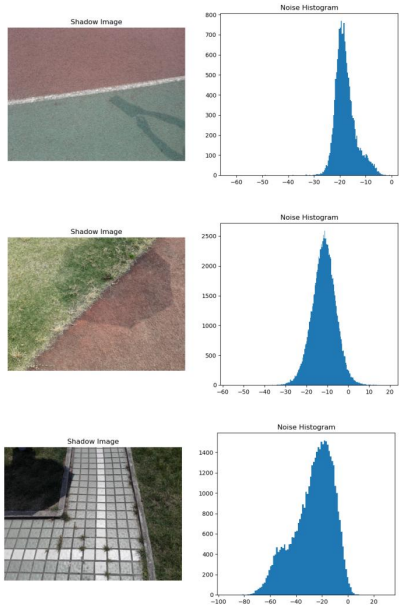

Fig. 2. Noise distribution computed from shadow image and shadow-free image.

3.4.2 Update Rule

In the image formation, we view the image as a vector. However, for shadow removal especially with a shadow mask, we only have to update the value of the shadow are. For efficient and convenient implementation, we input the whole shadow image and use it for all computation. For $\mathbf{z}$, $\mathbf{x}_{comp}$, $\boldsymbol{\upsilon}$, we only update the value in the shadow area and keep other values unchanged in each iteration (for example, we keep $\mathbf{Dx} - \mathbf{z} = 0$ in all non-shadow area).

3.4.3 Tuning of Parameters

In $\mathbf{z}$-update, we adjust the value of $\boldsymbol{\kappa}$, instead of setting it as constant in the whole process. We set $\boldsymbol{\kappa}$ based on $\boldsymbol{\upsilon}$ to give a new threshold.

$$\boldsymbol{\kappa} = \overline{|\boldsymbol{\upsilon}|} \tag{18}$$

3.5 Evaluation Metrics

We use Peak signal-to-noise ratio (PSNR), Structural similarity index (SSIM) and Root mean squared error (RMSE) as quantitative evaluation metrics for shadow removal. The comparison will be carried between the true shadow-free image and the output of our algorithm.

$$MSE = \frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n}[I_0(i,j) - I(i,j)]^2 \tag{19}$$

$$RMSE = \sqrt{MSE} \tag{20}$$

$$PSNR = 20 \cdot \log_{10}\left(\frac{MAX_{I_0}}{MSE}\right) \tag{21}$$

$$SSIM = \frac{(2\mu_I\mu_{I_0} + c_1)(2\sigma_{II_0} + c_2)}{(\mu_I^2 + \mu_{I_0}^2 + c_1)(\sigma_I^2 + \sigma_{I_0}^2 + c_2)} \tag{22}$$

where μ and σ^2 represent the mean pixel values and variances of the ground-truth shadow-free image (I_0) and the output (I), σ_{II_0} is the covariance, and c_1, c_2 are two variables to stabilize the division with weak denominator.

To get a desired and acceptable result, we expect lower RMSE and higher PSNR, SSIM.

4 Experimental Results

This section discusses the design of our experiments and then analyzes and evaluates the qualitative and quantitative results.

4.1 Dataset

The Image Shadow Triplets dataset (ISTD) is a dataset for shadow understanding that contains 1870 image triplets of shadow image, shadow mask, and shadow-free image [10] under 135 scenes (Fig. 3). We choose it for our experiments since it gives explicit shadow masks, which help specify shadow area, and shadow-free images, which can be used for comparison in evaluation. Furthermore, since we propose a simple framework and it doesn't require training, all the triplets, which contain various categories of shadows in different scenes, can be used for assessing the performance of the algorithm. The input will be a duplet containig a shadow image and the shadow mask. The corresponding shadow-free (ground-truth) image will be used in the evaluation stage for comparison with the output of the algorithm.

4.2 Shadow Removal Results

We feed the shadow image and image mask as input and get a "shadow-free" image as output with properly-set parameters. The shadow removal results are shown in Fig. 4. The original shadow images are shown in the first column and the corresponding results got from the algorithm are shown in the second column.

As can be seen in the figure, the algorithm works well and gives expected results, with shadow being removed and the shadow area restored. For the experimental images, $\rho = 0.08$ is applied.

4.3 Tuning of ρ

Here, we want to figure out the impact of ρ. In Equation 12, ρ determines the weight of the penalty term $\|\mathbf{Dx} - \mathbf{z}\|^2$. We experiment with several values of ρ to see how it impacts the restoration of the image (Fig. 5).

According to all evaluation metrics (Fig. 6), the performance of the algorithm becomes poorer as ρ increases. And with ρ decreasing below a specific and small threshold (e.g. 0.08), the results don't differ from each other on a large scale. We can see from Fig. 5 that when $\rho = 20$, an evident edge-like artifact appears in the image and this can be explained by the significantly weighted penalty term.

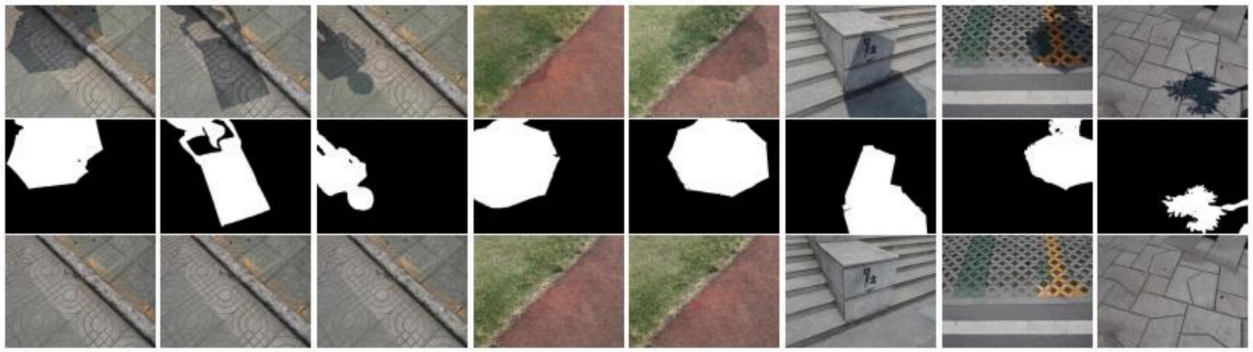

Fig. 3. An illustration of several shadow, shadow mask and shadow-free image triplets in ISTD [10].

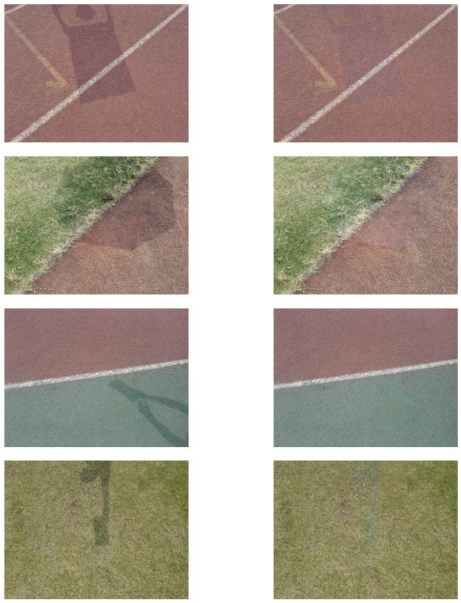

Fig. 4. Shadow removal results. Column (1) input shadow image. (2) output shadow-free image.

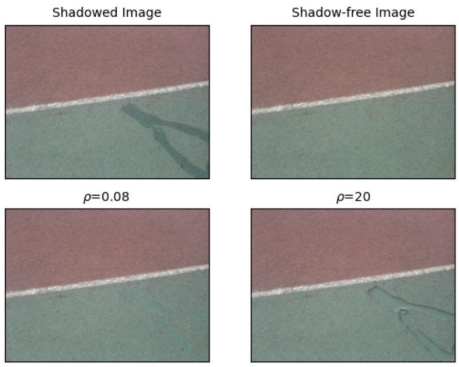

Fig. 5. Different ρ for shadow removal.

TABLE 1
Evaluation for Images in Fig. 7

	Hard 1	Hard 2	Hard 3	Soft
PSNR	21.54	18.50	22.66	33.32
SSIM	0.94	0.89	0.85	0.95
RMSE	0.084	0.118	0.074	0.021

4.4 Results for Different Categories of Shadows

As we mentioned in Section 3.2, there are different kinds of shadow (Fig. 7) and we try to experiment with them to test the adaptability of the algorithm.

The soft shadows retain the texture of the background surface, whereas the hard shadows are too dark and have little texture [9]. For shadow areas with more textures, the removal becomes more difficult and the result is not as good as expected. The artifacts concentrate on obvious shadow edges, insufficient removal, discrepant chroma matching and loss of textures. As shown in TABLE 1, the worst values for each metric are marked in red. For the second image, the output is least satisfactory visually and has the lowest PSNR and highest RMSE. For the third image, the SSIM is the lowest, which quite makes sense as most of the shadows appear in the area with most and complicated textures.

5 Discussion

The advantages of the algorithm include the simplicity of the framework, efficient and convenient processing of the images, time-saving feature. However, it also has its disadvantages, such as poor adaptability to various categories of shadows, insufficient restoration in chroma and textures. Apart from these, parameter tuning is required for each input image and this is not a patch on the learning-based methods.

There are several limitations on the algorithm. Firstly, we assume that the shadows are Gaussian-distributed. However, the realistic illumination condition is unknown and can be complicated. Large amount of sample images are inconsistent with the assumption. Then, we implement the removal process only in the Y (luminance) channel. The change in Y can affect the whole image and the

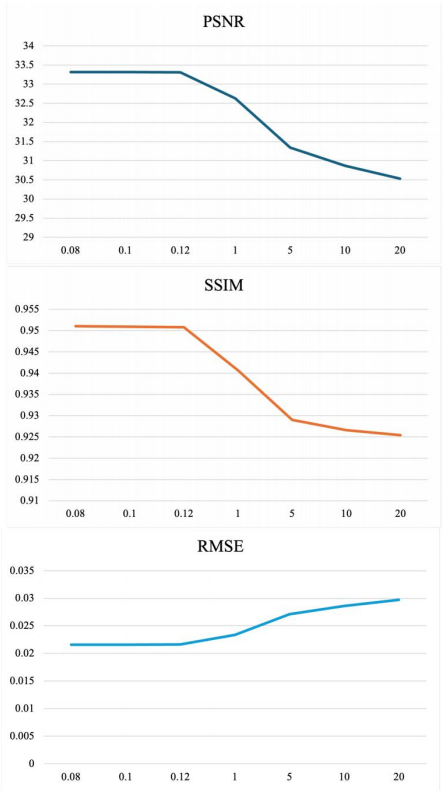

Fig. 6. PSNR, SSIM, RMSE values for different ρ.

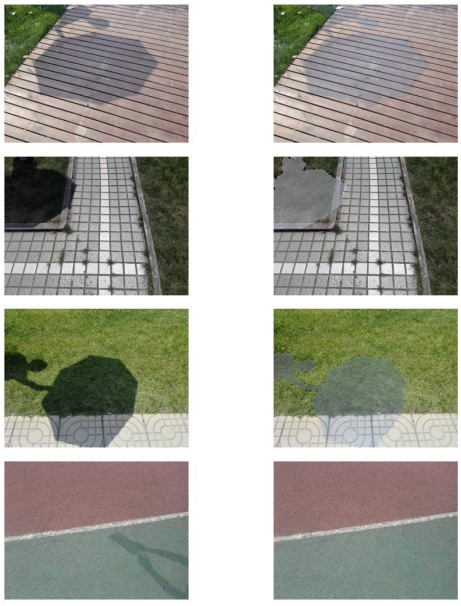

Fig. 7. Shadow removal results for hard and soft shadows. Column (1) input shadow image. (2) output image.

chroma channels visually. Lastly, the parameters are human-determined and indicates a loss of generalization of the algorithm.

For these limitations, here are several possible improvements. We can try different distributions to model the shadow and take the illumination information into account. Then, we should consider adjustment in the chroma channels and try to retain the textures. Lastly, we can combine the idea with neural networks training and increase the adaptability of the framework.

6 Conclusion

We present a framework for shadow removal from images based on an inverse problem formulation and the Half-Quadratic Splitting (HQS) method. The proposed approach aims to provide an efficient and convenient solution for enhancing image quality by removing shadows. The performance of the algorithm is evaluated using three metrics, PSNR, SSIM, RMSE. The results demonstrate the effectiveness of the proposed method in removing shadows and improving image quality, as evidenced by the quantitative metrics.

Despite its promising results, there are several potential areas for improvement and future research. Incorporating additional priors or constraints related to shadow formation and illumination models could further enhance the accuracy of the shadow removal process. Additionally, exploring more advanced optimization techniques or deep learning-based approaches may lead to better performance and generalization capabilities.

In conclusion, the proposed framework offers a practical and accessible solution for shadow removal.

Acknowledgments

The author would like to thank Gordon Wetzstein, Qingqing Zhao, Axel Levy and Cindy Nguyen for their advice on this project and the original work and code that this shadow removal project is based on.

References

[1] S. M. Pizer, E. P. Amburn, J. D. Austin, R. Cromartie, A. Geselowitz, T. Greer, B. ter Haar Romeny, J. B. Zimmerman, and K. Zuiderveld, "Adaptive histogram equalization and its variations," *Computer vision, graphics, and image processing*, vol. 39, no. 3, pp. 355–368, 1987.

[2] E. H. Land, "The retinex theory of color vision," *Scientific american*, vol. 237, no. 6, pp. 108–129, 1977.

[3] G. D. Finlayson, S. D. Hordley, and M. S. Drew, "Removing shadows from images using retinex," in *Color and imaging conference*, vol. 2002, no. 1. Society for Imaging Science and Technology, 2002, pp. 73–79.

[4] S. H. Khan, M. Bennamoun, F. Sohel, and R. Togneri, "Automatic shadow detection and removal from a single image," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 3, pp. 431–446, 2016.

[5] D. J. Ketcham, "Real-time image enhancement techniques," in *Image processing*, vol. 74. SPIE, 1976, pp. 120–125.

[6] S. M. Pizer, "Intensity mappings for the display of medical images," *Functional mapping of organ systems and other computer topics*, pp. 205–217, 1981.

[7] R. Hummel, "Image enhancement by histogram transformation," *Unknown*, 1975.

[8] D. Geman and C. Yang, "Nonlinear image recovery with half-quadratic regularization," *IEEE transactions on Image Processing*, vol. 4, no. 7, pp. 932–946, 1995.

[9] K. Deb and A. H. Suny, "Shadow detection and removal based on ycbcr color space," *SmartCR*, vol. 4, no. 1, pp. 23–33, 2014.

[10] J. Wang, X. Li, and J. Yang, "Stacked conditional generative adversarial networks for jointly learning shadow detection and shadow removal," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 1788–1797.