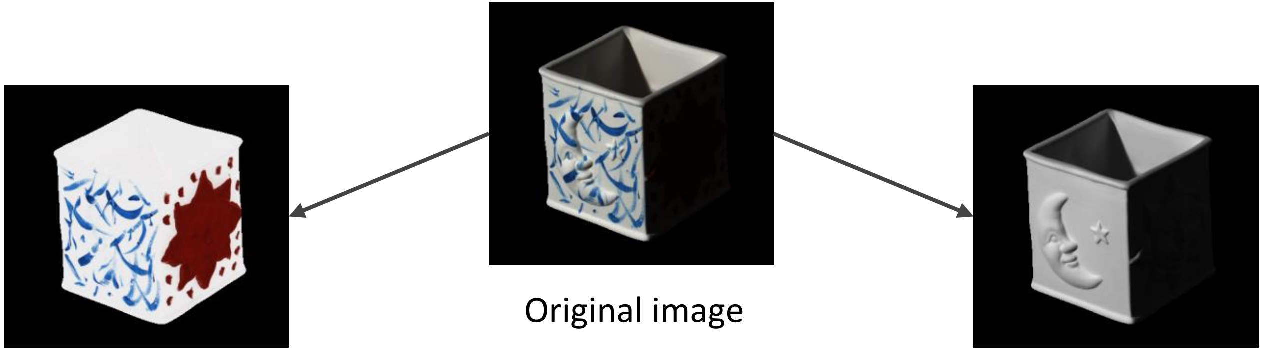
2021 IEEE International Conference on Image Processing (ICIP)
19 - 22 September 2021 Anchorage, Alaska, USA

AN OPTICAL PHYSICS INSPIRED CNN APPROACH FOR INTRINSIC IMAGE DECOMPOSITION

Harshana Weligampola, Gihan Jayatilaka, Suren Sritharan (*presenter*)
Parakrama Ekanayake, Roshan Ragel, Vijitha Herath, and Roshan Godaliyadda.

Intrinsic Image Decomposition (IID)

Intrinsic Image Decomposition (IID) is the problem of **decomposing an image into its constituents**



Original image

Reflectance

Shading

IID - Background

- Supervised IID
  - Impractical due to the absence of large datasets with ground truth.

- Unsupervised IID
  - Not robust enough to decompose images with various scene types in different lighting conditions.
  - Techniques based on human vision system (e.g. retinex theory) do not exploit the existing physics understanding of light into account in improving the image decomposition.

Proposed Physics based IID

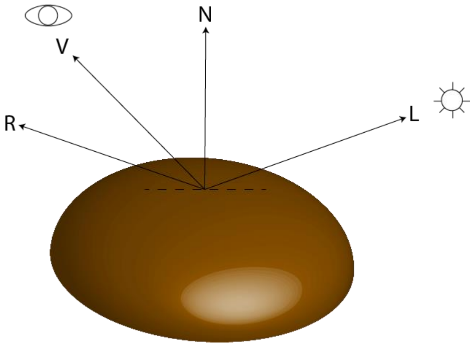
A novel **unsupervised IID** framework inspired by **optical physics** capable of decomposing a **wide variety** of images captured under different lighting conditions.

Image Interpretation - Phong's model

The perceived pixel intensity of an image is given by the **Phong's model**.

The main components of an image are:

1) Ambient light intensity
2) Diffused light intensity
3) Specular light intensity



Phong's model - Mathematical Formulation

The pixel Intensity at point p is,

$$I_p = \int_\lambda \underbrace{k_a r_p(\lambda) i_a(\lambda)}_{\text{Ambient}} + \sum_{\hat{L}^{(n)} \in \mathbf{L}} \{\underbrace{k_d r_p(\lambda)[\hat{L}^{(n)}.\hat{N}_p] i_d^{(n)}(\lambda)}_{\text{Diffused}} + \underbrace{k_s s_p(\lambda)[\hat{R}^{(n)}.\hat{V}]^\gamma i_s^{(n)}(\lambda)}_{\text{Specular}}\} d\lambda$$

Ambient Diffused Specular

Phong's model - Mathematical Formulation

First consider a narrow band (λ_c),

$$I_p(\lambda_c) = k_a r_p(\lambda_c) i_a(\lambda_c) + \sum_{\hat{L}^{(n)} \in \mathbf{L}} \{k_d r_p(\lambda_c)[\hat{L}^{(n)}.\hat{N}_p] i_d^{(n)}(\lambda_c) + k_s s_p(\lambda_c)[\hat{R}^{(n)}.\hat{V}]^{\gamma} i_s^{(n)}(\lambda_c)\}$$

Then assuming :

1) The ambient illumination is constant
2) Only one light source exists
3) Specular term is negligible
4) Ignoring coefficient

$$I_p(\lambda_c) = \underbrace{r_p(\lambda_c)}_{\text{Reflectance}} \quad \underbrace{[\hat{L}.\hat{N}_p] i_d(\lambda_c)}_{\text{Shading}}$$

Parameters derived from the image

1. **Reflectance Ratio Gradient (RRG)**

   Identify the boundaries of the uniform reflectance in an image.

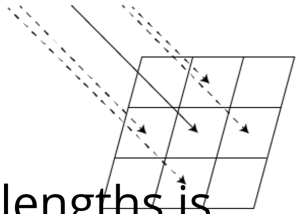
2. **Reflectance Approximation Map (RAM)**

   Approximate reflectance of an image.

3. **Shading Gradient (SG)**

   Gradient of the shading map of an image.

Reflectance Ratio Gradient



The **log ratio** of pixel intensity between 2 narrow band wavelengths is,

$$\mathcal{J}_p(\lambda_a, \lambda_b) = log\left(\frac{I_p(\lambda_a)}{I_p(\lambda_b)}\right) = log\left(\frac{r_p(\lambda_a) i_d(\lambda_a)}{r_p(\lambda_b) i_d(\lambda_b)}\right)$$

Consider the gradient

Assuming that single wavelength intensity of adjacent pixels are equal

$$\nabla \mathcal{J}(\lambda_a, \lambda_b) = \nabla log\left(\frac{r(\lambda_a)}{r(\lambda_b)}\right)$$

RRG

Reflectance Approximation Map

If $I_p(\lambda_a) = I_p(\lambda_b)$,then the log ratio $\mathcal{J}_p(\lambda_a, \lambda_b) = 0$.

If $I_p(\lambda_a) \gg I_p(\lambda_b)$,then the log ratio $\mathcal{J}_p(\lambda_a, \lambda_b)$ will be high and vise-versa

Based on this the RAM is used to approximate the reflectance (*R*) as,

$$\mathrm{RAM}_R = \frac{\overline{\mathcal{J}}_p(\lambda_R, \lambda_G) + \overline{\mathcal{J}}_p(\lambda_R, \lambda_B)}{2}$$

RAM : Red channel

where $\bar{\mathcal{J}}_p(\lambda_R, \lambda_G)$ is the value clipped in [0, 1]

The RAM highlights the significant wavelength in the reflectance (*R*)

Shading Gradient Map

Consider $\mathcal{K}_p(\lambda_c) = \log\left(I_p(\lambda_c)\right)$. If 2 neighboring pixels have the same reflectance, then the gradient is independent of reflectance and illumination.

$$\nabla\mathcal{K}(\lambda_c) = \mathcal{K}_{p_1}(\lambda_c) - \mathcal{K}_{p_2}(\lambda_c) = \nabla log\left([\hat{L}.\hat{N}]\right)$$

In such points, the gradient is independent or reflectance and illumination.
So, SG is defined as,

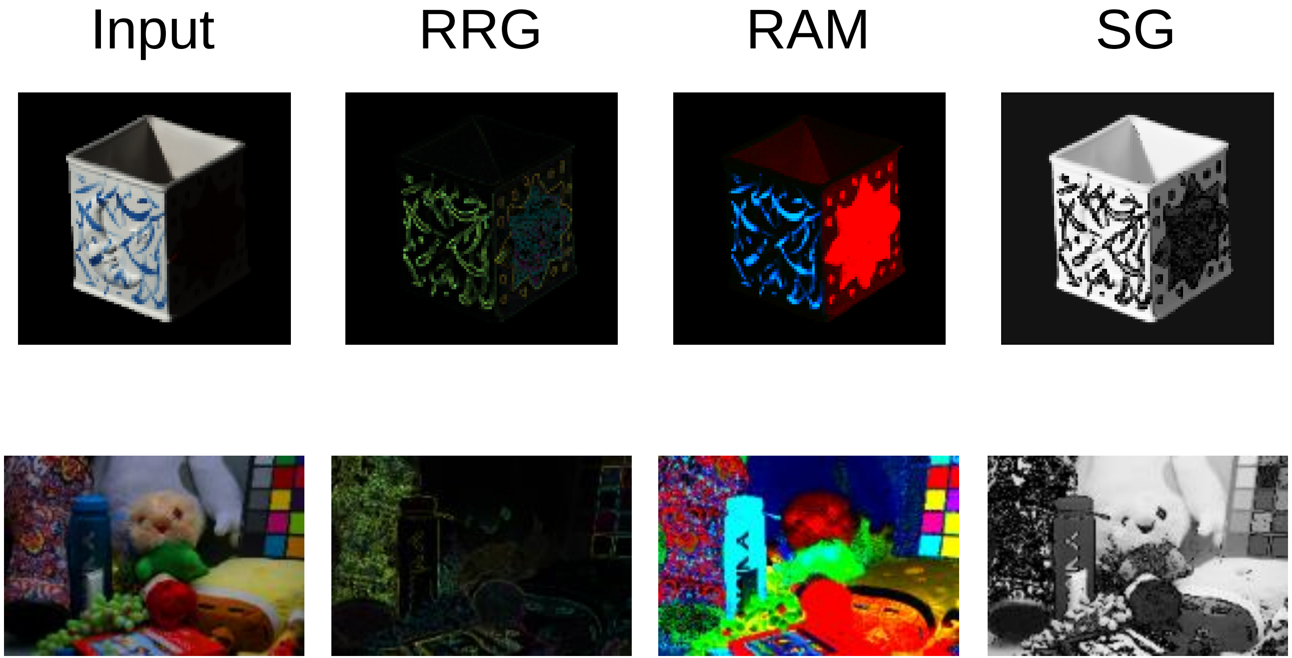
$$\nabla\mathcal{K}(\lambda_c) = \begin{cases} \nabla log\left([\hat{L}.\hat{N}]\right) & \text{if } M_{RRG}^{(c)} < 0.1 \\ 0 & \text{otherwise} \end{cases} \quad ; c \in \{R, G, B\}$$

SG Map

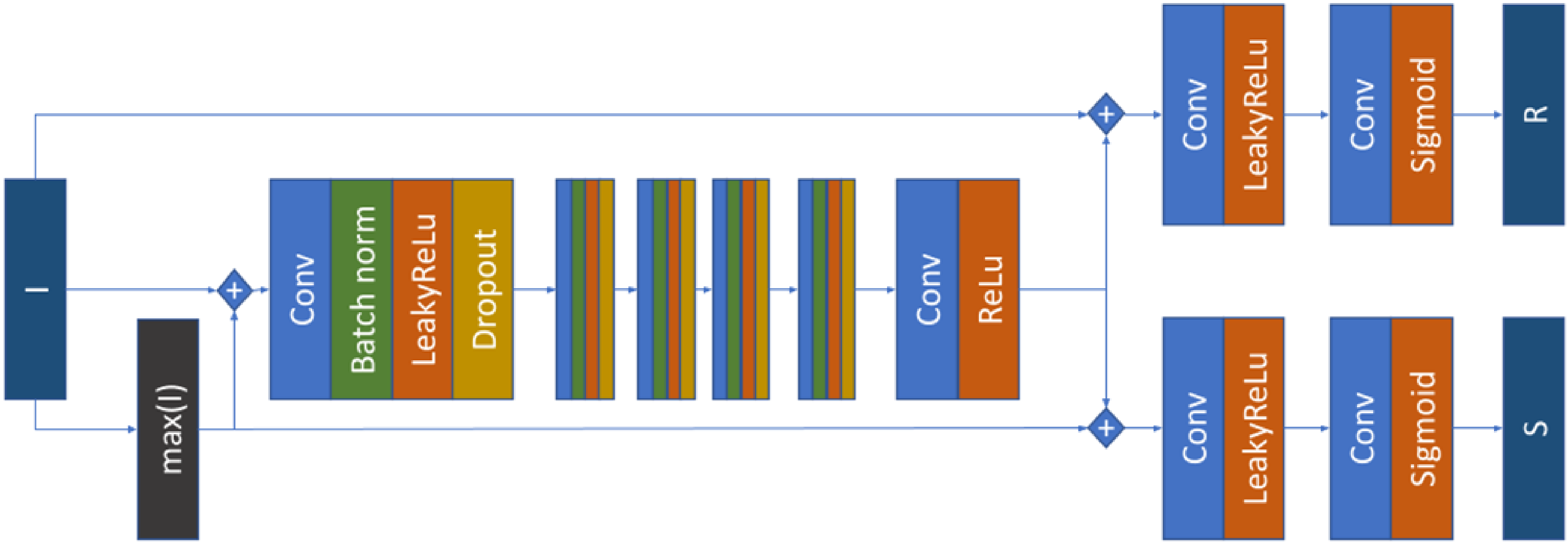
where $M_{RRG}^{(c)}$ is a filter to find pixels with same reflectance as their neighbors.

$$M_{RRG}^{(R)} = \frac{\nabla\mathcal{J}(\lambda_R, \lambda_G) + \nabla\mathcal{J}(\lambda_R, \lambda_B)}{2}$$

RRG, RAM and SG images



Network Architecture



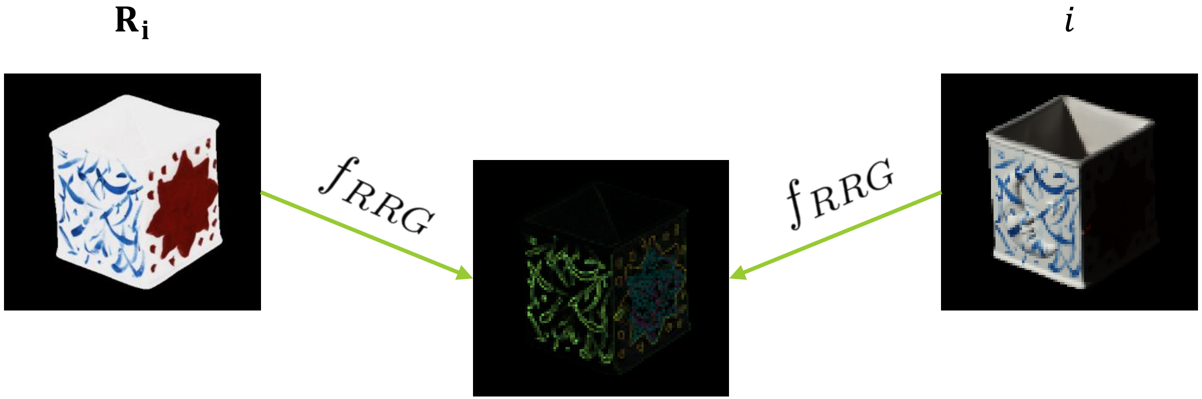
Proposed loss function

$$\mathcal{L} = \alpha_1 \mathcal{L}_{recon} + \alpha_2 \mathcal{L}_{ss} + \alpha_3 \mathcal{L}_{rrg} + \alpha_4 \mathcal{L}_{sg} + \alpha_5 \mathcal{L}_{ram}$$

1. Reconstruction loss
2. Shading smoothness loss
3. Reflectance Ratio Gradient (RRG) loss
4. Reflectance Approximation Map (RAM) loss
5. Shading Gradient (SG) loss

Reflectance Ratio Gradient Loss (1/5)

Ensures that the RRG of reflectance (*R*) is similar to the RRG of the image (*i*)

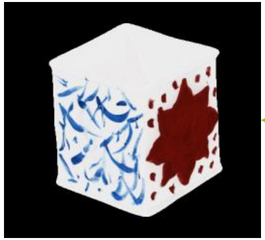


$$\mathcal{L}_{rrg} = ||f_{RRG}(\mathbf{R}_i) - f_{RRG}(i)||_1$$

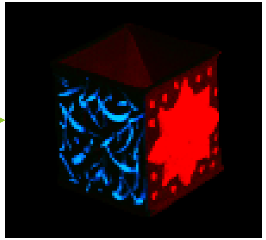
Reflectance Approximation Map Loss (2/5)

Ensures that the "**significant wavelength**" in the image and the reflectance are the same.

$\mathbf{R_i}$



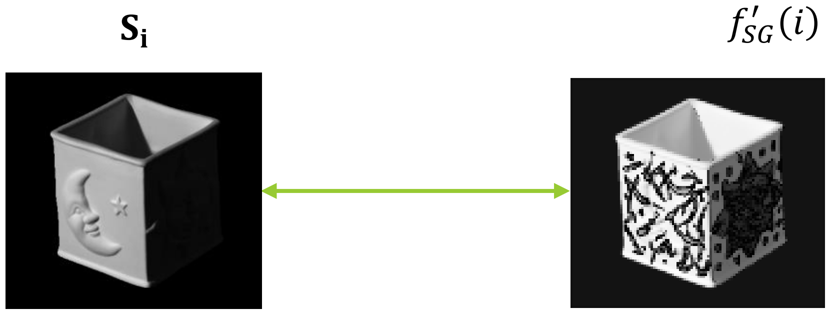
$f_{RAM}(i)$



$$\mathcal{L}_{ram} = ||(\mathbf{R}_i - f_{RAM}(i)) \times f_{RAM}(i)||_1$$

Shading Gradient Loss (3/5)

Ensures that the shading gradient of the image is similar to the natural logarithmic gradient of the shading component.



$$\mathcal{L}_{sg} = ||(\nabla \log(\mathbf{S_i}) - f'_{SG}(i)) \times f'_{SG}(i)||_1$$

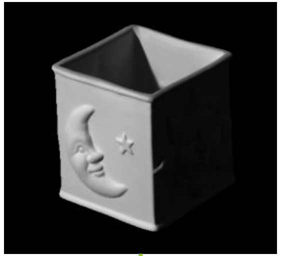
Reconstruction Loss (4/5)

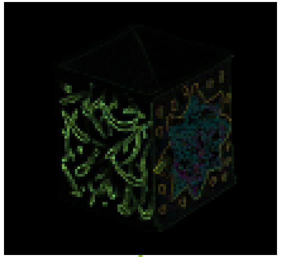
Ensures that the image (I) can be reconstructed from the Reflectance (R) and Shading (S).
(assumption: all reflectance maps are invariable of the lighting condition)

$$\mathcal{L}_{recon} = ||\mathbf{R}_i\mathbf{S}_i - \mathbf{I}_i||_1$$

Shading Smoothness Loss (5/5)

Ensures that the shading map (S) is smooth where the RRG is smooth.





$$\mathcal{L}_{ss} = ||\nabla \mathbf{S}_i \exp(-10 f_{RRG}(i))||_1$$

Training, Validation and Testing

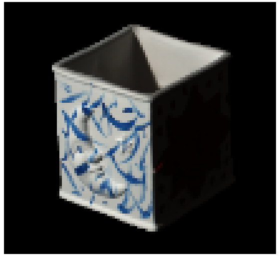
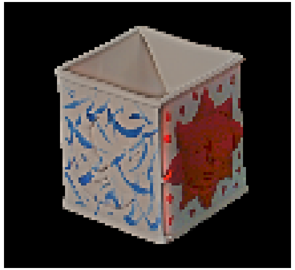
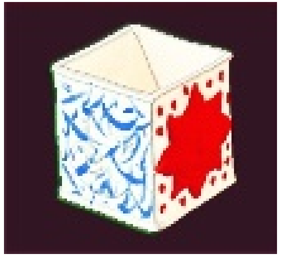
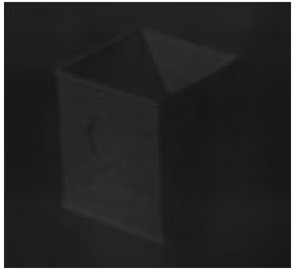
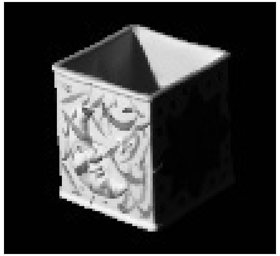
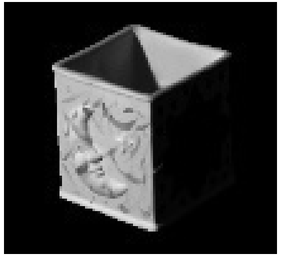
- Training and Validation:
  - Dataset : LOL dataset. (485 training, 15 validation)
  - Epochs : 100 epochs using Adam optimizer.
  - Learning Rate : 0.002 initial with a exponential decay factor of 0.01

- Testing
  - Dataset : MIT dataset (20 images)

Visual comparison of output (1/3)

Input	Letry et.al	CGIntrinsic	RetinexNet	Ours	
					Reflectance
					Shading

Visual comparison of output (2/3)

Input	Letry et.al	CGIntrinsic	RetinexNet	Ours	
					Reflectance
					Shading

Visual comparison of output (3/3)

Input	Letry et.al	CGIntrinsic	RetinexNet	Ours	
					Reflectance
					Shading

Numerical comparison - Validation Set

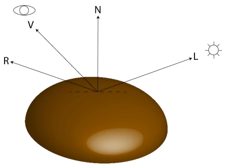
Metric / Method	LOL : 15 images			
	RMSE ↓	PSNR ↑	SSIM ↑	NIQE ↑
Letry et.al	21.87	<u>35.28</u>	**0.96**	7.55
CGIntrinsic	63.28	18.95	0.36	**14.78**
Retinex-net	<u>6.88</u>	34.64	0.90	<u>7.63</u>
Ours	**2.00**	**43.12**	<u>0.95</u>	<u>7.63</u>

↓ : Lower is better

Numerical comparison - Test Set

Metric / Method	MIT : 20 Images				MIT (R)		MIT (S)	
	RMSE	PSNR	SSIM	NIQE	RMSE	PSNR	RMSE	PSNR
Letry et.al	6.67	39.26	**0.99**	12.06	**41.91**	**16.58**	40.88	16.46
CGIntrinsic	40.95	17.36	0.11	**17.47**	48.47	16.28	59.62	12.99
Retinex-net	3.77	37.85	0.95	14.02	67.39	13.48	37.97	18.54
Ours	**1.04**	**41.66**	0.96	14.02	45.90	15.82	**30.54**	**20.14**

Summary

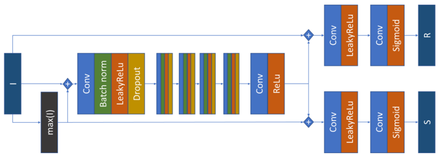


- Intrinsic Image Decomposition (IID) with simplified Phong's model.

$$\underbrace{I_p(\lambda_c)}_{\textbf{Image}} = \underbrace{r_p(\lambda_c)}_{\textbf{Reflectance}} \underbrace{[\hat{L}.\hat{N}_p] i_d(\lambda_c)}_{\textbf{Shading}}$$

- A set of maps (RRM, RAM, SG) to extract meaningful information from images.
- Optical physics inspired loss function and CNN model

$$\mathcal{L} = \alpha_1 \mathcal{L}_{recon} + \alpha_2 \mathcal{L}_{ss} + \alpha_3 \mathcal{L}_{rrg} + \alpha_4 \mathcal{L}_{sg} + \alpha_5 \mathcal{L}_{ram}$$



- Evaluating the model using numerical (RMSE, PSNR, SSIM, NIQE) and visual results.