Computer Vision Project Report

Low-Light Image Enhancement

Singam Setty Deshik

—

2101AI32

—

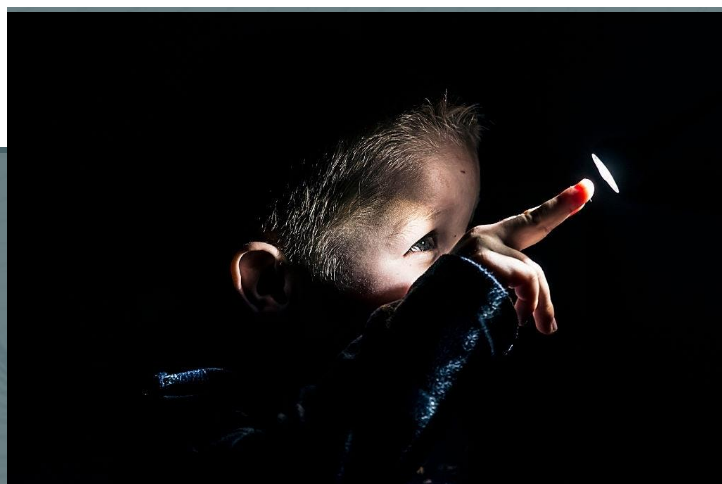
Professor Jimson Mathew

INTRODUCTION

When capturing images with a conventional digital camera in a back-lit environment, areas of the subject not directly exposed to light often appear excessively dark, resulting in images with reduced visibility. This phenomenon arises due to the limited dynamic range of the camera, which is narrower compared to that of the human eye.

This might affect viewers experience and lead to wrong message being communicated, such as inaccurate object/face recognition.

I have used three image enhancement techniques namely zero reference-DCE(deep curve estimation), single backlit image enhancement and a prompt learning for unsupervised backlit image enhancement.



METHODOLOGY

Zero-Reference Deep Curve Estimation (Zero-DCE)

This method formulates light enhancement as a task of image-specific curve estimation with a deep network. It trains a lightweight deep network, DCE-Net, to estimate pixel-wise and high-order curves for dynamic range adjustment of a given image. Zero-DCE is appealing in its relaxed assumption on reference images, i.e., it does not require any paired or unpaired data during training.

The key components in Zero-DCE, namely LE-curve, DCE-Net, and non-reference loss functions

Light-Enhancement Curve (LE-curve)

There are three objectives in the design of such a curve:

- each pixel value of the enhanced image should be in the normalized range of [0,1] to avoid information loss induced by overflow truncation;
- this curve should be monotonous to preserve the differences (contrast) of neighboring pixels; and
- the form of this curve should be as simple as possible and differentiable in the process of gradient backpropagation.

$$LE_n(\mathbf{x}) = LE_{n-1}(\mathbf{x}) + \mathcal{A}_n(\mathbf{x}) LE_{n-1}(\mathbf{x})(1 - LE_{n-1}(\mathbf{x}))$$

DCE-NET

To learn the mapping between an input image and its best-fitting curve parameter maps, we propose a Deep Curve Estimation Network (DCE-Net). We employ a plain CNN of seven convolutional layers with symmetrical concatenation.

Non-Reference Loss Functions

The following four types of losses are adopted to train our DCE-Net

1. Spatial Consistency Loss: The spatial consistency loss encourages spatial coherence of the enhanced image through preserving the difference of neighboring regions between the input image and its enhanced version
2. Exposure Control Loss: Exposure control loss, refers to the difference between the desired exposure level and the actual exposure level in an image
3. Color Constancy Loss: We design a color constancy loss to correct the potential color deviations in the enhanced image
4. Illumination Smoothness Loss: To preserve the monotonicity relations between neighboring pixels, we add an illumination smoothness loss to each curve parameter map

$$L_{total} = L_{spa} + L_{exp} + W_{col}L_{col} + W_{tv_A}L_{tv_A}$$

Single Backlit Image Enhancement

This algorithm proposes a simple and fast backlit image enhancement method that improves the visibility of dark parts. The proposed method uses gamma correction and histogram equalization to make an intensity-adjusted image and a contrast-enhanced image, respectively. Then, an alpha-blended image of the two images is obtained. Finally, an output image is obtained by fusing the original image and the

blended image using alpha blending to avoid causing artifacts in bright parts. The weight for fusion is calculated by using Otsu's method and the guided filter.

$$I_{ij} = (I_{ij}^{R} + I_{ij}^{G} + I_{ij}^{B})/3.$$

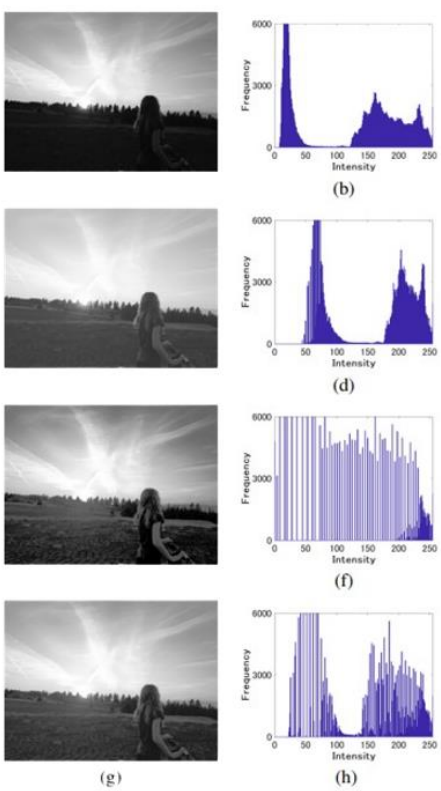
$$I_{ij}^{gamma} = 255 \cdot (I_{ij}/255)^{\frac{1}{\gamma}}.$$

$$O_{ij} = (1-\alpha) \cdot I_{ij}^{gamma} + \alpha \cdot I_{ij}^{HE},$$

$$W_{ij} = \begin{cases} 1, & I_{ij} < t \\ 0, & \text{otherwise} . \end{cases}$$

$$\tilde{O}_{ij} = \tilde{W}_{ij} \cdot O_{ij} + (1 - \tilde{W}_{ij}) \cdot I_{ij}.$$

$$\tilde{\boldsymbol{O}}_{ij}^{RGB} = \boldsymbol{I}_{ij}^{RGB} \cdot (\tilde{O}_{ij}/I_{ij}).$$



Processed backlit images and their histograms. Intensity image I.

Result of gamma correction for I.

Result of histogram equalization for I.

Enhanced image obtained by alpha blending.

Unsupervised backlit image enhancement method, abbreviated as CLIP-LIT, by exploring the potential of **Contrastive Language-Image Pre-Training (CLIP)** for pixel-level image enhancement.

The approach consists of two stages, In the first stage, we learn an initial prompt pair (negative/positive prompts referring to backlit/well-lit images) by constraining the text-image similarity between the prompt and the corresponding image in the CLIP embedding space.

With the initial prompt pair, used a frozen CLIP model to compute the text-image similarity between the prompts and the enhanced results to train the initial enhancement network.

In the second stage, refine the learnable prompts by utilizing backlit images, enhanced results, and well-lit images through rank learning

Evaluation Metrics:

PSNR (Peak Signal-to-Noise Ratio):

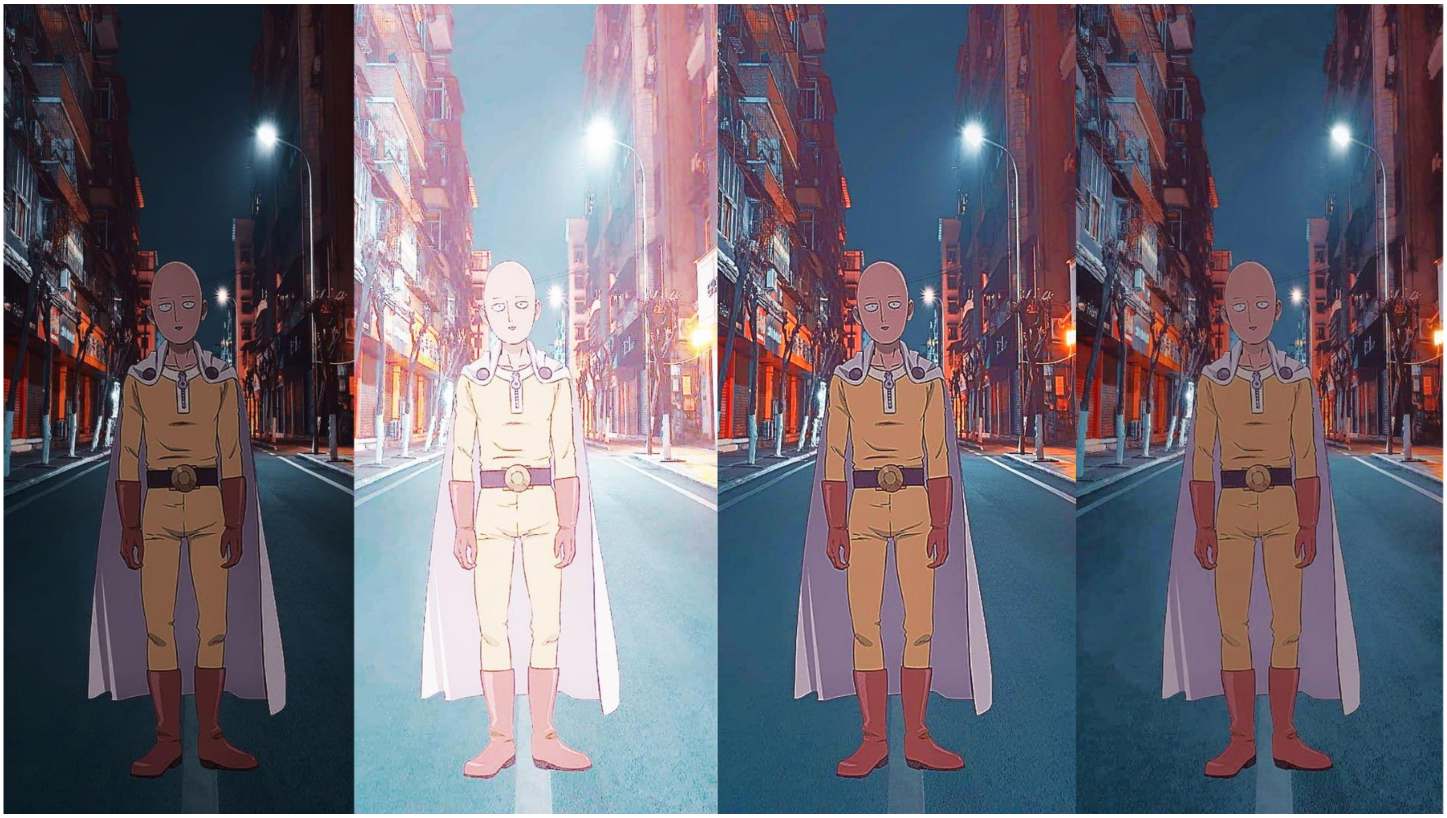
The Peak Signal-to-Noise Ratio (PSNR) is a widely-used metric in image processing and compression to quantify the quality of reconstructed or enhanced images by measuring the ratio of the peak signal power to the noise power. It provides a numerical measure of fidelity between the original and processed images, with higher PSNR values indicating better reconstruction quality.

NIQE (Natural Image Quality Evaluator):

The Natural Image Quality Evaluator (NIQE) is a perceptual image quality assessment metric designed to quantify the perceived quality of natural images. NIQE leverages statistical features derived from the image to estimate the overall quality, mimicking human perception. It operates without requiring reference images and has been demonstrated to correlate well with subjective human judgments of image quality.

Results

	Zero-DCE	Backlit Image Enhance	Prompt Learning(CLIP)
PSNR	5.71dB	22.76dB	15.60dB
NIQE	6.44	7.12	5.85



	Zero-DCE	Backlit Image Enhance	Prompt Learning(CLIP)
PSNR	4.68dB	17.12dB	13.96dB
NIQE	5.15	4.14	4.46

References:

https://keras.io/examples/vision
https://github.com/HumanChwan/single-backlit-image-enhancement
https://zhexinliang.github.io/CLIP_LIT_page/