DIP Proposal

Team4
r11943018 林子軒、r11943022 范詠為

Paper title
Single Image Haze Removal Using Dark Channel Prior

Motivation

Single Image Haze Removal is a computer vision technique that aims to remove haze or fog from a single input image. This technique can improve the quality of images, making them more useful for various applications, such as surveillance, aerial photography, and computer vision. For example, in surveillance, removing haze can improve the accuracy of object detection and recognition, allowing for better identification of people or vehicles in the scene. In aerial photography, haze removal can improve the clarity of the terrain, making it easier to detect changes or anomalies in the landscape. In computer vision, haze removal can improve the accuracy of algorithms that rely on visual information, such as image classification, segmentation, or tracking.

Moreover, the Dark Channel Prior approach is a popular and effective method for Single Image Haze Removal, and it has been widely used in various applications. By implementing this approach, we can leverage the existing research and codebase, and adapt it to specific use cases or datasets. Additionally, the implementation of this approach can be a useful exercise in computer vision and image processing, allowing us to gain practical experience in image enhancement techniques and optimization algorithms.

Problem definition

The problem of single image haze removal refers to the task of enhancing the visibility and appearance of a hazy or foggy image by removing the atmospheric haze or fog. The Dark Channel Prior (DCP) is a widely used method for haze removal, which assumes that the intensity of a pixel in a haze-free image is close to the minimum intensity in its local neighborhood in the hazy image. The problem of Single Image Haze Removal Using Dark Channel Prior involves estimating the transmission map, which describes the degree of haze in each pixel, and then using it to remove the haze and restore the original scene. The goal is to produce a visually pleasing and natural-looking haze-free image that enhances the details and colors of the scene while preserving its overall appearance.

Algorithm

The dark channel prior is a technique that relies on a common feature of outdoor images without haze. Specifically, in most areas of the image that are not sky, at least one color channel will contain some pixels with very low intensity values that are nearly zero. This means that the minimum intensity in these areas is also very low.

In computer vision and computer graphics, the model widely used to describe the formation of a hazy image is

$$\mathbf{I}(\mathbf{x}) = \mathbf{J}(\mathbf{x})t(\mathbf{x}) + \mathbf{A}(1 - t(\mathbf{x})),$$

where I is the observed intensity, J is the scene radiance, A is the global atmospheric light, and t is the medium transmission describing the portion of the light that is not scattered and reaches the camera.

We assume that the atmospheric light A is given and first normalize the haze imaging equation by A.

$$\frac{I^c(\mathbf{x})}{A^c} = t(\mathbf{x})\frac{J^c(\mathbf{x})}{A^c} + 1 - t(\mathbf{x}).$$

Then, we calculate the dark channel on both sides

$$\min_{\mathbf{y}\in\Omega(\mathbf{x})}\left(\min_c \frac{I^c(\mathbf{y})}{A^c}\right) = \tilde{t}(\mathbf{x}) \min_{\mathbf{y}\in\Omega(\mathbf{x})}\left(\min_c \frac{J^c(\mathbf{y})}{A^c}\right) + 1 - \tilde{t}(\mathbf{x}).$$

As the scene radiance J is a haze-free image, the dark channel of J is close to zero due to the dark channel prior

$$J^{\text{dark}}(\mathbf{x}) = \min_{\mathbf{y}\in\Omega(\mathbf{x})}\left(\min_c J^c(\mathbf{y})\right) = 0.$$

Therefore, we can estimate the transmission t simply by

$$\tilde{t}(\mathbf{x}) = 1 - \min_{\mathbf{y}\in\Omega(\mathbf{x})}\left(\min_c \frac{I^c(\mathbf{y})}{A^c}\right).$$

In practice, the presence of haze is a fundamental cue for humans to perceive depth. This phenomenon is called aerial perspective. If we remove the haze thoroughly, the image may seem unnatural and we may lose the feeling of depth. So, we can optionally keep a very small amount of haze for the distant objects by introducing a constant parameter.

$$\tilde{t}(\mathbf{x}) = 1 - \omega \min_{\mathbf{y}\in\Omega(\mathbf{x})}\left(\min_c \frac{I^c(\mathbf{y})}{A^c}\right).$$

With the atmospheric light and the transmission map, we can recover the scene radiance. However, the directly recovered scene radiance J is prone to noise because the transmission t(x) is close to zero. Therefore, we restrict the transmission t(x) by a lower bound t0. The final scene radiance J(x) is recovered by

$$\mathbf{J}(\mathbf{x}) = \frac{\mathbf{I}(\mathbf{x}) - \mathbf{A}}{\max(t(\mathbf{x}), t_0)} + \mathbf{A}.$$

Expected results

The expected result is a clear and visually pleasing image that accurately represents the scene without the distortion caused by the atmospheric haze or fog. Additionally, the output should be a haze-free image that looks natural and visually appealing to the human eye. The degree of haze removal can vary depending on the specific parameters and settings used in the algorithm, but the overall goal is to produce an image that accurately reflects the scene as it would appear in clear weather conditions.

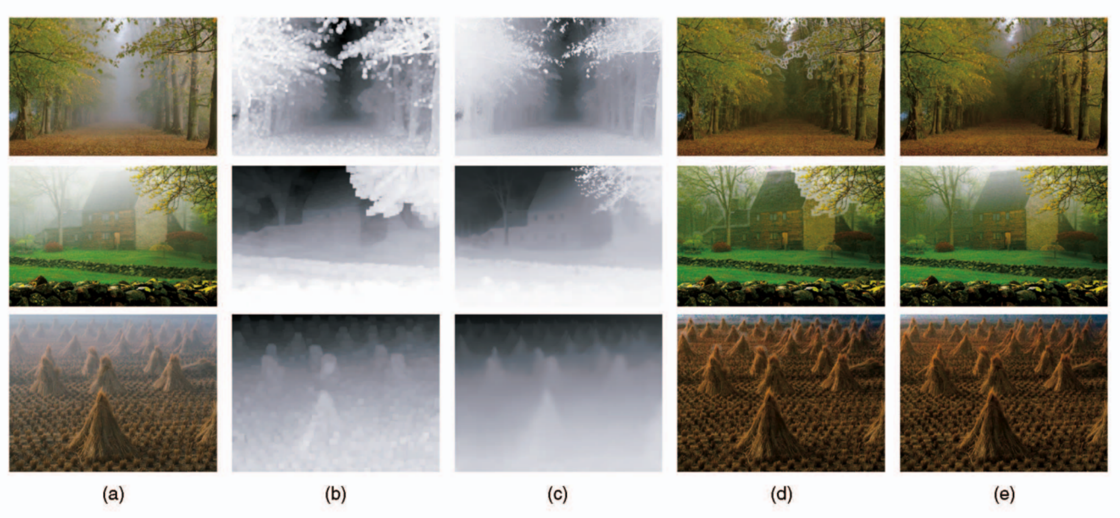

Fig. 6. Haze removal. (a) Input hazy images. (b) Estimated transmission maps before soft matting. (c) Refined transmission maps after soft matting. (d), (e) Recovered images using (b) and (c), respectively.

PSNR (Peak Signal-to-Noise Ratio) and SSIM (Structural Similarity Index Measure) are both metrics used to evaluate the quality of digital images or videos.

PSNR measures the quality of an image or video based on the ratio of the signal power to the noise power, relative to the maximum possible signal power (i.e., the peak signal). It is commonly used to compare the quality of two images or videos, especially in the field of image or video compression. A higher PSNR value indicates better quality.

SSIM measures the structural similarity between two images or videos, taking into account the luminance, contrast, and structural similarity of the images or videos. It aims to capture the perceptual quality of an image or video, rather than just the signal-to-noise ratio. A higher SSIM value indicates greater similarity between the two images or videos being compared, and thus higher quality.

PSNR can help us quantify the error between the restored image and the original image, and is often used as a loss function for optimization algorithms. SSIM, on the other hand, can consider the structure and perceptual quality of the image or video more comprehensively, and therefore can better reflect the visual quality of the image or video. Using both methods can provide a more comprehensive evaluation of the quality of the restored image.

Table 17 PSNR and SSIM comparison of existing techniques on RESIDE dataset

Method	SOTS Outdoor		SOTS Indoor	
	PSNR	SSIM	PSNR	SSIM
DCP [63]	19.13	0.82	16.62	0.82
AOD-Net [75]	20.29	0.88	19.06	0.85
DehazeNet [74]	22.46	0.85	21.14	0.85
GFN [77]	21.55	0.84	22.30	0.88
FFA-Net [86]	33.57	0.98	35.77	0.98
GMAN [78]	28.19	0.96	20.53	0.81
Deep DCP [88]	24.08	0.93	19.25	0.83
CAP [64]	22.30	0.91	19.05	0.84
MSCNN [73]	21.73	0.83	17.57	0.81
NLD [66]	18.07	0.80	17.29	0.75
BCCR [49]	15.49	0.78	16.88	0.79
Y-NET [23]	26.61	0.95	–	–
Deep Energy (Network) [92]	24.07	0.93	–	–
Improved CycleGAN [178]	21.78	0.80	–	–
GCANet [84]	–	–	30.23	0.98
HIDEGAN [185]	25.54	0.88	24.71	0.87
RYFNet [186]	–	–	21.44	0.87
DM^{2}F-Net [128]	–	–	34.29	0.98
DPDP-Net [141]	–	–	20.18	0.88

$$PSNR = 10\log_{10}\left(\frac{255^2}{MSE}\right) \quad SSIM(r,i) = \left(\frac{2\mu_r\mu_i + c_1}{\mu_r^2 + \mu_i^2 + c_1}\right)\left(\frac{2\mu_{ri} + c_2}{\sigma_r^2 + \sigma_i^2 + c_2}\right)$$

Reference

main algorithm

He, Kaiming, Jian Sun, and Xiaoou Tang. "Single image haze removal using dark channel prior." *IEEE transactions on pattern analysis and machine intelligence* 33.12 (2010): 2341-2353.

survey-type

Agrawal, Subhash Chand, and Anand Singh Jalal. "A comprehensive review on analysis and implementation of recent image dehazing methods." *Archives of Computational Methods in Engineering* 29.7 (2022): 4799-4850.