Digital Image Processing Lab Report 1

Image Dehazing

Sipei Li	Yuxuan Pang	Pei Yang	Jiuchuan Zhang
2019190505002	2019190505022	2019190505033	2019190505038

Abstract

An image may suffer from low contrast or poor resolution under hazy weather, and image dehazing is very important in image restoration. Current image dehazing techniques mainly consist of traditional approaches that utilise some fundamental relations or important observations and AI-based methods that depend on deep neural networks. In this lab report, we emphasised investigating traditional dehazing techniques. We benchmarked histogram equalisation, Dark Channel Prior, Single Scale Retinex, and Homomorphic filtering for image dehazing and Spatio-Temporal Markov Random Field (ST-MRF) for video dehazing on given datasets. While each method has its advantage in enhancing specific images, the Dark Channel Prior best qualitatively preserves unhazed images, and ST-MRF generated more natural results than the dehazed images at a faster speed per frame. Additionally, we proposed a novel method for generating dark channels that could restore the colour information better than the original Dark Channel Prior. Future works can focus on introducing quantitative metrics for dehazing performance evaluation. Section Heading

1 Introduction

As camera-equipped portable digital terminals have become popular in recent days, there is more chance for people to capture images or videos in the open air. However, due to the vagaries of climate and different illumination conditions, a picture could be degraded due to foggy weather or having excessively dark or bright regions. Up to now, there exist various technique, including intensity transformation that converts the intensity levels of an input image using a specific transformation function (linear or logarithmic) to convert the intensity levels; Histogram modelling, which enhance an input image by modifying image histogram, as a uniform, desired shape, and this is called histogram equalization; Homomorphic filtering is an image enhancement method that based on image formation model and represents input pixels with a product of illumination and the reflectance; In Jobson et al.'s Retinex method the illumination and reflectance are first estimated. Next, the estimated term is enhanced. Finally, the output image is obtained by merging the enhancement terms; And the Dark Prior channel by He et al [1].

2 Related Works

2.1 Dark Channel Prior for Image Dehazing

It is worth noticing that He et al.'s method is physically valid and can handle distant objects in heavily hazy images. They do not rely on a significant variance of transmission or surface shading, and their results have few halo artifacts.

As mentioned above, they also applied the Atmospheric scatting model, dehazing is to solve the J from I with a good evaluation of t, and A. J(x)t(x) is called direct attenuation, describe scene radiance and its decay in the medium; A(1-t(x)) is air light previously scattered light, and this leads to the shift of the scene colours. The author counted 5000 images taken in the open air and concluded that clear images' dark channel is concentrated between 0-15, and the dark channel prior is defined as

$$J^{dark}\ x\ = \min_{y\in\Omega(x)} (\min\big(J^c\ y\ \big)),$$

where J^c is the colour channel of J, and $\Omega\ x$ is the local patch centre at x.

By doing a Two-sided equivalent transformation on the atmospheric scatting model

$$\min_{y\in\Omega\ x}\left(\min_{c}\frac{I^c\ y}{A^c}\right) = t\ x\ \min_{y\in\Omega\ x}\left(\min_{c}\frac{J^c\ y}{A^c}\right).$$

Since the dark channel always approaches zero, and A^c is positive, we can estimate $t\ x$.

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

To keep the feeling of depth and the photorealism parameter, ω is introduced.

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

To estimate A, since the sky has an infinity distance between light and camera, $t\ x$ is close to 0, from the model $I\ k\ \approx A$, however this may result in the algorithm finds the brightest object instead of the area where the haze concentrated most instead of this. [1] chose 0.1% brightest point in channel I^{dark}, and then corresponding this point to I to determine if they are suitable for A's estimation. By applying this method, the result is robust.

After we get the estimation for t and A, we get

$$J = \frac{I - A}{t} + A.$$

2.2 Spatial-Temporal Markov Random Field for Video Dehazing

Video dehazing has a wide range of real-time applications, and these challenges mainly come from Spatio-temporal consistency and computational efficiency. [2] claimed three main assessment aspects for video dehazing:

Spatial consistency. There are two constraints on spatial consistency. The property that fog concentration is locally invariant is used to overcome the estimation noise. Furthermore, his recovered video should be as natural as the original video to handle internal frame discontinuities—temporal coherence. The human visual system is inconsistently sensitive to time. However, simply using the still image dehazing algorithm frame by frame may break the temporal coherence of the video, and the recovered video has severe springing artifacts—computational efficiency. The algorithm must be able to efficiently process a large number of pixels in a video sequence. In particular, a practical real-time dehazing method should achieve a speed of more than 15 frames per second.

In their model, they applied the Atmospheric scattering model given by

$$I\ \ x\ \ = J\ \ x\ \ T\ \ x\ \ + A(1 - T\ \ x\ \ ),$$

where x represents pixel indexes, $I\ \ x$ is the observed image, $J\ \ x$ is the real scene to be discovered, A is the global atmospheric light that remains constant in the whole image, and $T\ \ x$ is the transmission rate for light reaches the camera without scattering. For medium transmission estimation, Cai et al. [2] suggested $\omega = 0.7$.

In most haze-free images, at least one-color channel has some pixels whose intensity values are very low and close to zero. The dark channel is defined as the minimum channel in RGB color space; For Spatial-Temporal MRF: This is caused by the air light; the intensity value is increased while haze concentration is enhanced. The spatial likelihood function is given by

$$P_s\ \ w, b\ \ \propto \prod_{y \in \Omega\ \ x} \exp\left(-\frac{\|w\ \ x\ \ V\ \ y\ \ + b\ \ x\ \ - D\ \ y\ \ \|_2^2}{\sigma_s^2}\right),$$

where $\Omega\ \ x$ is a local patch centred at x with size $r \times r$, and σ_s is the spatial parameter. Temporal Coherence temporal MRF is used for temporal coherence, and at time t, its likelihood function is defined by

$$P_\tau\ \ w_t, b_t\ \ \propto \prod_{\tau \in [-f, +f]} \exp\left(-\frac{\|w_t\ \ x\ \ V_t\ \ x\ \ + b_t\ \ x\ \ - D_{t+\tau}\ \ x\ \ \|_2^2}{\sigma_\tau^2}\right).$$

[2]'s method can alleviate fluctuations and reduce flickering artifacts more efficiently. Compared with the other methods, [2]'s work prevents image over-saturation from happening and keeps spatio-temporal coherence. They achieved 116.371fps for CIF 352×288 on Intel i7 3770 CPU. However, it did not perform well in estimating the haze concentration with high accuracy.

3 Experiments: Histogram Computation (Task 1 and 2)

Figure 1 plots two hazed images and one clear image with their corresponding grayscale histograms.

The first image is low contrast due to its pixel brightness clusters around 100 to 175. Although the pixel values spread more widely across different intensity levels in the second image, they still cluster around two values. The differentiation of the two clusters could result from the high contrast between the foreground leaves and the background city landscape. At the same time, the close brightness levels within each group correspond to the low contrast within the foreground or the background region. Finally, the third image has a clear landscape with a more equally-distributed histogram.

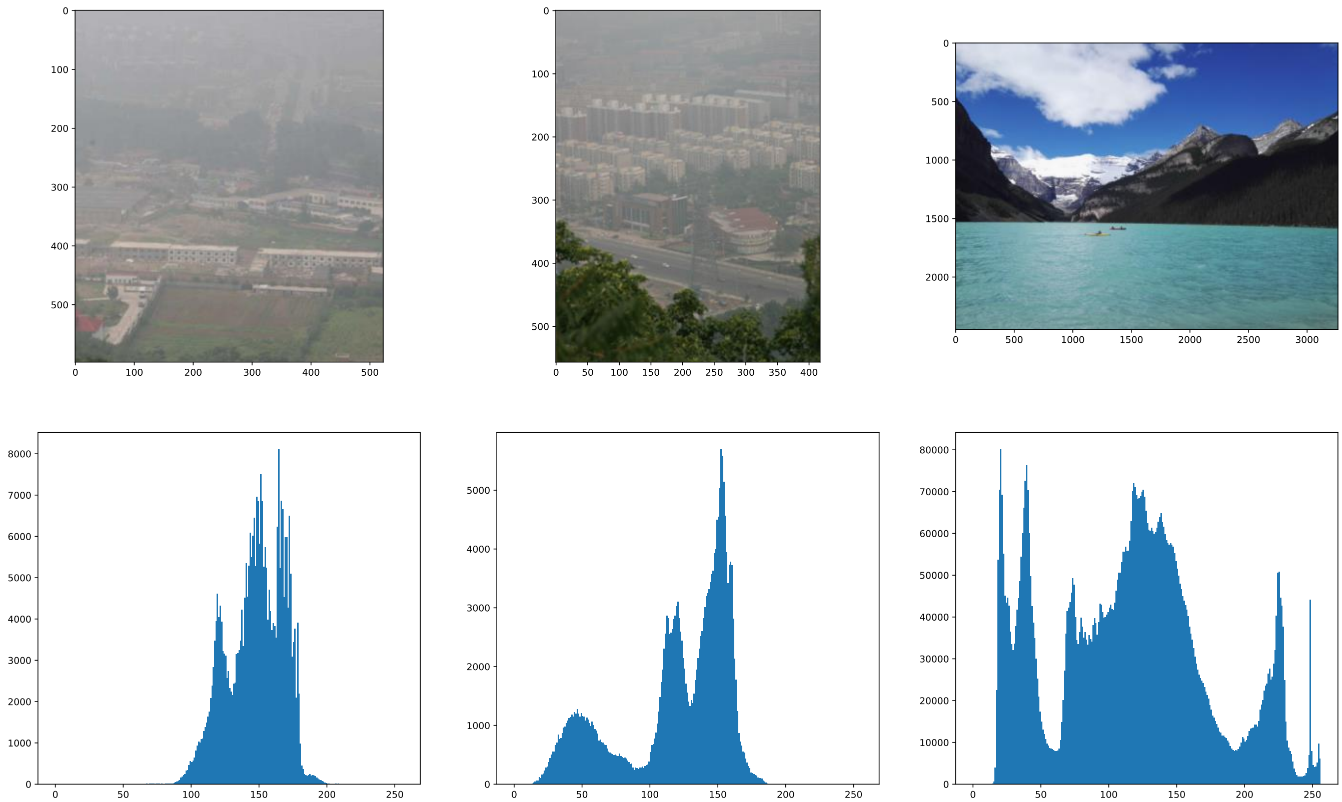


Figure 1. Two hazed images and one clear image with their corresponding grayscale histograms.

4 Experiments: Image Dehazing (Task 3.1)

4.1 Experimental Setup

We performed the experiments on the provided IEI2019 dataset. This dataset includes 22 hazed images and two clear images. On each image, we tested the following listed dehazing techniques.

(a) Histogram equalisation. This refers to first converting an original image from its own colour space to the HSV domain and then perform histogram equalisation on the value channel on the original image. The image will then be converted back to its original colour space.

(b) Single Scale Retinex (SSR). This refers to applying Choi et al.'s proposed method [3] for dehazing.

Figure 2-1. Comparison of different dehazing techniques on two hazed images (a) and (b) and two clear images (c) and (d), including (1) original image, (2) histogram equalisation, (3) Single Scale Retinex (SSR), (4) Homomorphic filtering on grayscale original images, (5) the previous technique followed by a histogram equalisation, (6) Homomorphic filtering on the value channel of the original image in the HSV colour space, (7) Dark Channel Prior, (8) Dark Channel Prior followed by a histogram equalisation, and (9) our proposed method. Our proposed method is good at restoring the colour information while dehazing, although it may not recover as much details compared to the original Dark Channel Prior.

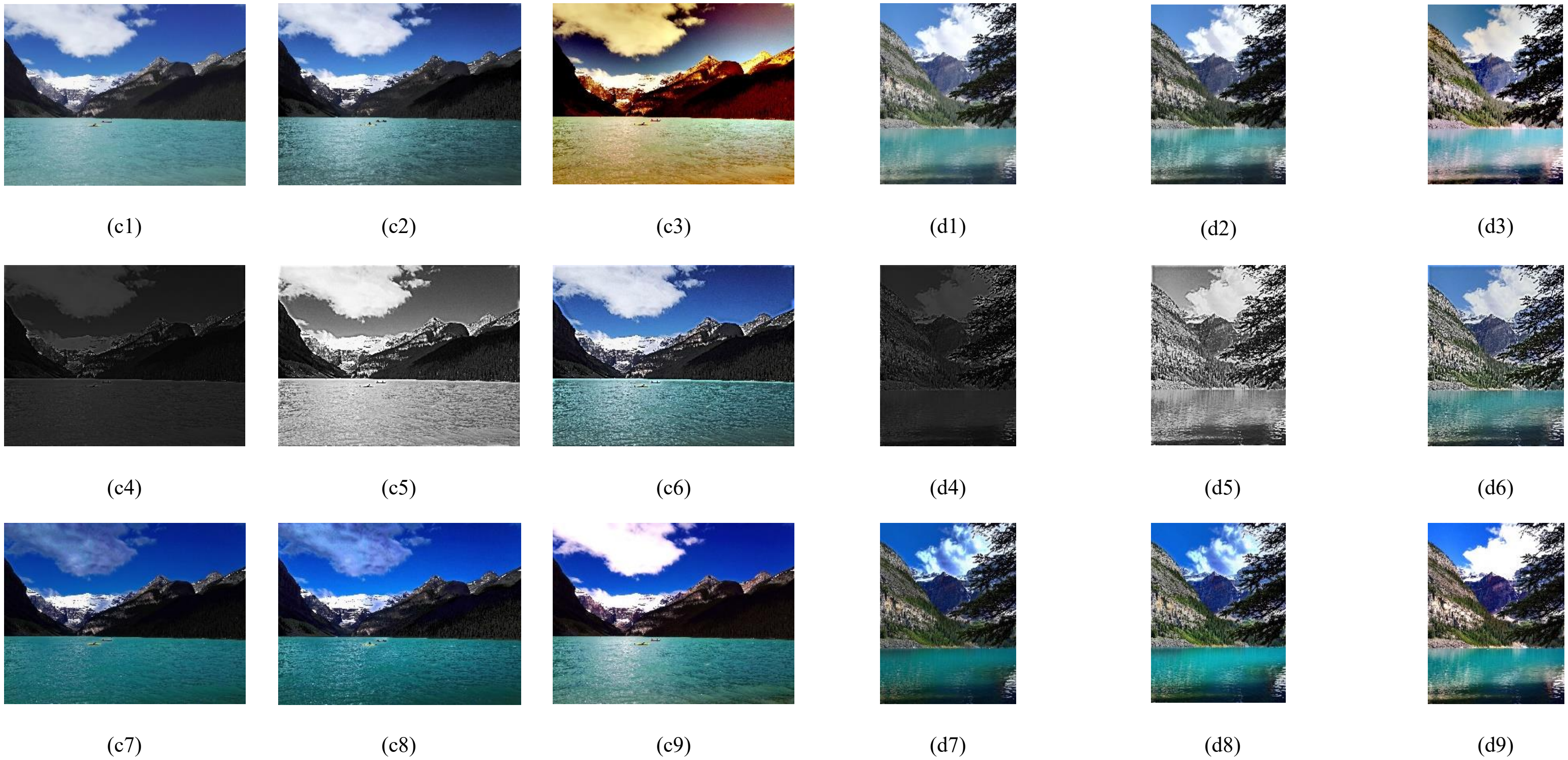


Figure 2-2. Comparison of different dehazing techniques on two hazed images (a) and (b) and two clear images (c) and (d), including (1) original image, (2) histogram equalisation, (3) Single Scale Retinex (SSR), (4) Homomorphic filtering on grayscale original images, (5) the previous technique followed by a histogram equalisation, (6) Homomorphic filtering on the value channel of the original image in the HSV colour space, (7) Dark Channel Prior, (8) Dark Channel Prior followed by a histogram equalisation, and (9) our proposed method. Our proposed method is good at restoring the colour information while dehazing, although it may not recover as much details compared to the original Dark Channel Prior.

Figure 3. (a) The grayscale version of an original image. (b) Histogram equalised version of the grayscale image. (c) Dark channel generated by our proposed method. (d) The original dark channel. It can be observed that (c) and (d) are similar apart from the dilation in (d), and the main difference in this case is an offset.

(c) Homomorphic filtering on grayscale images. Since homomorphic filtering should be performed on a two-dimensional image, this approach refers to performing homomorphic filtering on the grayscale version of a hazed image.

(d) Homomorphic filtering on grayscale images followed by histogram equalisation. This is a cascade of (e) with a histogram equalisation performed on this image.

(e) Value channel homomorphic filtering. This refers to converting the original image first into the HSV colour space and then perform homomorphic filtering on the value channel.

(f) Value channel homomorphic filtering followed by histogram equalisation. This is a cascade of (g) with (a). For simplicity, the two colour space conversions between (g) and (a) can be cancelled out.

(g) Dark Channel Prior. This refers to applying He et al.'s proposed work in [1].

(h) Dark Channel Prior followed by histogram equalisation. This is equivalent to a cascade of (a) and (b).

(i) Our proposed method. This is a modified version of Dark Channel Prior [1] with a different dark channel generation algorithm as in the following equation. Our algorithm is given by

$$c = |x - x_{hist}|,$$

where c is the generated dark channel, x is the grayscale version of the original image. Performing histogram equalisation on x yields x_{hist}. Our proposal was based on an empirical observation that this method often distinguishes the hazed regions with unhazed ones. Figure 3 gives one example of this empirical observation.

4.2 Qualitative Analysis on Dehazing Performance

Figure 2-1 compares the dehazing techniques on two hazed images, and Figure 2-2 compares them on two unhazed images. We will now qualitative evaluate each the dehazing performance of each method.

4.2.1 Histogram Equalisation

Figure 2-1 (a2) and (b2) shows that histogram equalisation is effective in increasing the contrast of the images. However, in Figure 2-1 (a2), the haze at the far end of the road is still not removed because the pixels there are originally similar in value. Another downside of histogram equalisation is that it may cause the lose of details, for example, the tree branches at the top in Fig 2-1 (a2).

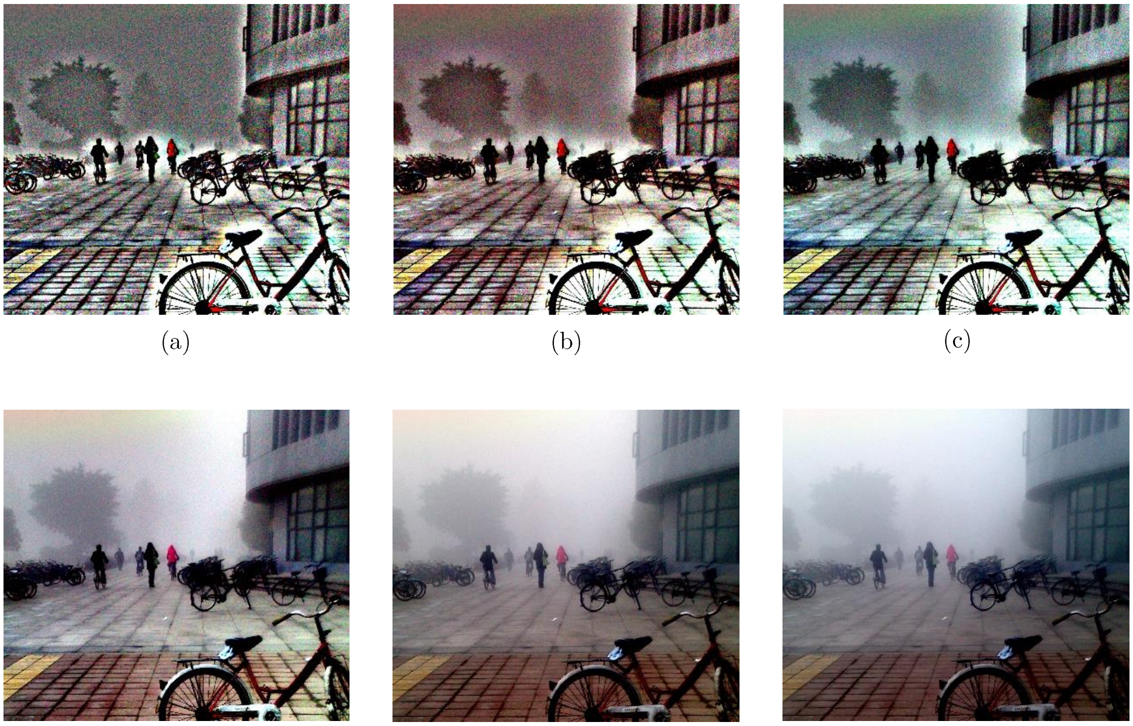
Additionally, histogram equalisation on the value channel generates similar dehazed images as before, which is possibly due to no specific efforts were made to restore the colour information. Overall, on hazed images, histogram equalisation is good at removing the haze in regions that suffers less from haze, but it could worsen the situation in regions that are heavily hazy.

On clear images, there are possibilities for histogram equalisation to erase details, for example, the illuminated side of the mountains in Figure 2-2 (b2).

4.2.2 Single Scale Retinex

Unlike histogram equalisation, Figure 2-1 (a3) and (b3) shows that SSR is good at removing haze and revealing details in regions that are heavily hazed. Comparing among all the images, Figure 2-1 (a3) best reveals the details at the far end of the road.

However, the images generated by SSR suffers from relatively poor fidelity compared to the ideal unhazed scene. In Figure 2-1 (a3), the ground surface was added with perturbation patterns and even colours. An additional disadvantage is that SSR requires manually adjust the variance, and Figure 4 illustrates that an inappropriate variance may lead to significant performance degradation in dehazing. This indicates that SSR is not suitable for image batch processing.



(a) (b) (c)

(d) (e) (f)

Figure 4. Dehazing performance of SSR with variance equals (a) 30, (b) 60, (c) 90, (d) 300, (e) 600, and (f) 900. At low variances, SSR generates a considerable amount of noise perturbations. At very high variances, the performance of dehazing can hardly be observed. Therefore, selecting an appropriate variance for an image before SSR processing is crucial.

On the clear landscape image in Figure 2-2 (c3), SSR significantly influenced the colour in Figure 2-2 (a2), which indicates this method cannot be generally applied for image batch processing. This again proves that SSR is not suitable for image batch processing.

4.2.3 Homomorphic Filtering

Figure 2-1 (a4) and (b4) both show that homomorphic filtering on grayscale images generates very dark images, which directly leads to two shortcomings. One, the output images are too dark to be visually inspected. A very close pixel values may let an image suffer heavily from losing details due to quantisation.

To visually inspect the output images, a straight-forward approach would be applying either a contrast stretching or a histogram equalisation. Figure 2-1 (a5) and (b5) shows the output after performing histogram equalisation. The equalised Figure 2-1 (a5) shows that Homomorphic filtering cascaded with a histogram equalisation outperforms a stand-alone histogram equalisation in that it both sharpens the haze tree branches and reveals the details at the far end of the road more clearly. On Figure 2-1 (b5), homomorphic filtering obtained a good performance, revealing many high-frequency details on the terminal and the plane at the far end. However, like SSR, it also deteriorates the clear part of the image such as the ground surface is the previous image.

Switching the Homomorphic filtering from the grayscale image to the value channel of an HSV image, only limited colour information is restored, and the output images are still very similar to their grayscale counterparts unless with careful visual inspection. This could be due to no efforts have been conducted on restoring the colour-related information (hue and saturation).

Homomorphic filtering sharpens details on clean images like Figure 2-2 (c4-6) and (d4-6) show very strong amplification on mild perturbations, which could be considered as strong deteriorations of the contour information from the original image. The images are also considered to have more uniform illuminations than their original counterparts. Its negative performance on clean images indicate that Homomorphic filtering, in general, is not appropriate for automatic image processing.

4.2.4 Dark Channel Prior

Dark Channel Prior generates very dark dehazed images but with rich detail. In Figure 2-1 (a7), the top branches from the trees were very well preserved. Unlike homomorphic filtering, cascading histogram equalisation with Dark Channel Prior will lose many details, including the tree branches at the top or the aircraft livery in Figure 2-1 (b7). Alternatively, applying homomorphic filtering may significantly deteriorate the colour fidelity of an original image. For example, the saturation in Figure 2-1 (a8) is abnormally high, and the region of the sky above the main road became magenta.

Dark Channel Prior preserves the original images well, apart from losing some details like the tree contours in the forest in the middle right of Figure 2-2 (c7). Applying histogram equalisation to the result yields restoration of the tree contours but disappearing of stones as sparsely distributed dark pixels on the snow mountain. However, compared to the above analysed methods, the dark-channel-based methods generally causes much imperceptible deteriorations on clean images.

4.2.5 Our Proposed Method

Our proposed method, similar to histogram equalisation and SSR, may cause the image to lose details. However, compared to the above-analysed methods, our proposed method, by visual inspection, obtains more natural colours after image restoration. For instance, in Figure 2-1 (b9), the taxiways were brown, which is the only warm colour in images from (b2) to (b9).

On clean images, our proposed method deteriorates the original images slightly more perceptible than the original Dark Channel Prior in causing overexposures. However, compared to other methods analysed above, such deteriorations can be considered as mild.

4.3 Quantitative Analysis on Time Consumption

Table 1 summarises the per image time consumption of the methods. The processing speed of most dehazing techniques on the dataset was, on average, between 0.5 to 1 second per image. Among the compared methods, histogram equalisation was significantly faster than all other methods, at around

0.01 seconds per image, which makes it promising in batch processing. This is possibly due to that histogram equalisation is a statistical-based technique. Oppositely, the time consumption of SSR grows exponentially with increasing variance to more than two minutes when the variance equals 900.

Table 1. Average time consumption of different dehazing techniques on the dataset.

Methods	**Average Time Consumption (seconds per image)**
EqualiseHist	0.01
DarkChannel	0.63
DarkChannel+EqualiseHist	0.97
SSR@3	0.83
SSR@6	0.7
SSR@9	0.51
SSR@15	3.14
SSR@30	6.01
SSR@60	9.61
SSR@90	18.9
SSR@150	42.98
SSR@300	89.95
SSR@900	121.26
HomoFilter-Gray	0.43
HomoFilter-Gray+EqualiseHist	0.98
HomoFilter-HSV	0.45
HomoFilter-HSV+EqualiseHist	0.46

5 Experiments: Video Dehazing (Task 3.2)

5.1 Experimental Setup

We conducted the experiments on four videos on the IEV2022 dataset: airplane, Huangshan, and Tian'anmen. Additionally, we dehazed a 29-second 4K video footage shot in front of the main building of the UESTC campus. We evaluated Spatial-Temporal Markov Random Field (ST-MRF) for video dehazing.

5.2 Qualitative Analysis on Dehazing Performance

Figure 5 shows the dehazing performance of ST-MRI on a footage in IEV2022 recorded from a cable car at the Yellow Mountain, and Figure 6 shows the performance on another aerial footage shot at the UESTC campus.

(a) (b)

Figure 5. Dehazing performance of ST-MRF on a video in IEV2022 shot at the Yellow Mountain. Images in group (a) are frames from the hazed videos, and group (b) shows their dehazed counterparts. Images in group (b) shows higher contrast compared to image (a), as the illumination appears to be more uniform. Both frames in group (b) show very natural restored colour without unreasonably high saturations or semantically unexpected colours.

Dehazing performed on both videos significantly improved the contrast of the image frames. Unlike contrast stretching, ST-MRF improves the contrast more on hazy (or cloudy) regions more than a region that is comparatively unhazed. This yields a more equally distributed illumination. In all videos processed by ST-MRF, this illumination is consistently distributed across different frames due to its spatial-temporal processing, leading to smoother and more reasonable dehazing results.

(a) (b)

Image 6. Dehazing performance of ST-MRF on an aerial video shot at the UESTC campus in Chengdu. Images in group (a) are frames from the hazed videos, and group (b) shows their dehazed counterparts. It is clearly shown that in image (b), apart from revealing more details (contours) of the buildings around the horizon, ST-MRF also restored the colour information by increasing the overall saturation – for example, the sky is bluer after dehazing.

Especially, the dehazing outcome of the UESTC aerial footage provides a promising result of the ST-MRF. The frames in Figure 6(a) were shot in a hazy day, with the buildings around the horizon not clearly visible, and the difference in colour of the buildings difficult to distinguish by human eyes. However, in the dehazed frames in Figure 6(b), the sky became bluer with the far-shot buildings with very high-saturation colours.

5.3 Quantitative Analysis on Processing Speed

The ST-MRC processed the 4K (3840×2160) UESTC aerial footage at 2.4264fps, or 49.688ns per pixel on an AMD Ryzen 5 5600X processor at 4.50GHz. This is shorter than the single-image dehazing techniques investigated in Section 3 by at least an order of magnitude on the same processor.

6 Conclusions

We tested several traditional image and video dehazing techniques in this lab, including histogram-, Dark Channel-, Retinex-, and Homomorphic filtering-based approaches. For image dehazing, all compared methods display some unique advantages and shortcomings, and combining some of the methods could be a promising option. On clean images, however, while Dark Channel Prior causes nearly imperceptible deteriorations, other techniques generally result in significant perturbations. For video dehazing, ST-MRF generated better frames both qualitatively in the dehazing performance and quantitatively in the processing speed. But still, the overall conclusion is that there is no best way to dehaze an image or a video, and the selection of dehazing techniques should be based on discretion on detailed task specifications.

References

[1] K. He, J. Sun, and X. Tang, “Single image haze removal using dark channel prior,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 12, pp. 2341–2353, 2011, doi: 10.1109/TPAMI.2010.168.

[2] [B. Cai, X. Xu, and D. Tao, “Real-time video dehazing based on spatio-temporal MRF,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 9917 LNCS, no. 201605030011, pp. 306–314, 2016, doi: 10.1007/978-3-319-48896-7_31.

[3] D. H. Choi, I. H. Jang, M. H. Kim, and N. C. Kim, “Color image enhancement using single-scale retinex based on an improved image formation model,” *Eur. Signal Process. Conf.*, no. Eusipco, 2008.