Revealing halos concealed by cirrus clouds

Yuji Ayatsuka
independent researcher: Tokyo, Japan

Correspondence: Yuji Ayatsuka (ayatsuka@acm.org)

Received: 12 March 2023 – Discussion started: 1 September 2023
Revised: 11 April 2024 – Accepted: 9 May 2024 – Published: 24 June 2024

Abstract. Many types of halos (including arcs) appear in the sky. Each type of halo corresponds to the shape and orientation of ice crystals in clouds, and reflects the state of the atmosphere; therefore observing them from the ground greatly helps in understanding the state of the atmosphere. However, halos are easily obscured by the contrast of the cloud itself, making it difficult to observe them. This difficulty can be overcome by enhancing halos in images, for which various techniques have been developed. This study describes the construction of a sky-color model for halos and a new effective algorithm to reveal halos in images.

1 Introduction

Ice crystals that form clouds (mainly cirri and cirrostratus) sometimes show rings and arcs of light (often colored) around the sun or the moon, as shown in Fig. 1. These optical phenomena are called halos.¹ Halos can be of many types, with each type corresponding to the shape and orientation of ice crystals in clouds (Greenler, 1980). Halos provide valuable information about the atmosphere. Observing halos from the ground is an important way to understand meteorological processes. It is a wide-area observation of ice crystals in clouds. For example, the difference in frequency of appearance between 22 and 46° halos suggests the ratio of pristine to non-pristine crystals in clouds (van Diedenhoven, 2014). There are also several studies of ice crystals and halo observations (Lynch and Schwartz, 1985; Sassen et al., 1994; Um and McFarquhar, 2015; Sassen, 1980; Lawson et al., 2006).

An attempt is also being made to achieve automated observations of halos using a sun-tracking camera (Forster et al., 2017) or a total sky imager (Boyd et al., 2019). Image processing techniques can be used to observe even faint halos or other atmospheric optical phenomena in photographs more clearly, which can greatly aid in these types of studies (Riihonen et al., 2000; Moilanen and Gritsevich, 2022; Großmann et al., 2011).

A halo is formed by the interaction of sunlight with the particles in a cloud. However, the contrast of the cloud itself prevents us from seeing halos clearly. To resolve this difficulty, halo observers have developed various image processing techniques that enhance halos in images and improve their observation, and a $B - R$ (blue minus red) processing proposed by Rossetto (2011) and Lefaudeux (2013) is one such technique and is known to be an effective method for enhancing halos. It involves subtracting a pixel value in the red channel from that in the blue channel and setting the difference as the gray value of the pixel. As a result, the areas in an image that are more reddish or bluish than the surrounding areas are enhanced. The author of the present study has developed a revised method, called autoBR, which was implemented in an image processing tool, named Atmospheric Optical Image Enhancer (Ayatsuka, 2022). In autoBR, the red and blue channels are differently weighted and the green channel is also referenced (details are described in Sect. 3).

In contrast, in $B - R$ processing, the red and blue channels are differently weighted and the green channel is also referenced (details are described in Sect. 3). In contrast, in $B - R$ processing, the red and blue channels are equally weighted and the information in the green channel is not used.

Figure 2 shows a halo image processed with both $B - R$ processing and autoBR. It is quite noticeable that the contrasts of the clouds are flattened by the two processes, especially by autoBR. In other words, these processes are thought to flatten the “cloud colors” in the image and extract the ap-
Pearl and intensity of halos. It is quite useful for observing and analyzing halos precisely through ground images.

This study constructs and validates a sky-color model for halo display. Using the model, a new algorithm is developed, called sky-color regression, to more effectively cancel cloud colors and enhance halos in the image.

The rest of the paper is organized as follows. Section 2 discusses related work on image processing for similar purposes. Section 3 presents details of existing algorithms, $B - R$ processing, and autoBR. Section 4 proposes a sky-color model and validates it with an image of a halo display. Section 5 describes the new algorithm called sky-color regression. Section 6 discusses some remaining issues related to the new algorithm, and Sect. 7 concludes the paper.

2 Related work

As described in the previous section, the contrasts of clouds should be reduced to extract halos in an image. Similar processes have been studied for other purposes, for example, for dehaze or fog reduction. Cloud detection using color information obtained from images is another related research topic, although with an opposite purpose.

2.1 Dehaze/fog reduction

The main goal of dehaze/fog reduction is to compensate the poor camera image quality when observing land and sea scenes under hazy conditions (Chengtao et al., 2015; Singh et al., 2016). Building physical models of the hazed view to recover background information is one of the approaches that is similar to the one used in the present study. Shi et al. (2022) used blue and red channel information to improve images in sandstorms.

The main difference between these studies and the present one is that, in their target situation of the previous studies, complex background images are obscured by smooth haze or fog, whereas in the target situation of the present study, relatively simple patterns are obscured by clouds. Because the present study has a different purpose, which is to process the image for easier halo detection and not to compensate for the image quality, a simpler model than those described in previous studies can be employed.

Gao and Li (2017) described the removal of cirrus clouds from the ground images taken by a Landsat satellite by using a microwave imager. An image taken by a band that is highly reflected by cirrus clouds helps in effectively removing cirrus clouds from the corresponding visible light (RGB) image.

2.2 Cloud detection

Many studies have investigated the detection of cloud regions or segmentation in sky images taken from the ground for meteorological observation. Koehler et al. (1991) set a threshold on the ratio of red and blue pixels that can be utilized to judge whether a pixel is a part of clouds.

Dev et al. (2017) investigated the contribution of composite parameters for a cloud segmentation algorithm, including values in color spaces other than RGB, such as HSV and $L^*a^*b^*$. They found that the saturation in HSV; the ratio of red and blue in RGB; and $\frac{B - R}{R + B}$, also called the “normalized blue / red ratio” in Li et al. (2011), are important for cloud...
Figure 2. Image results obtained using of the $B - R$ processing and autoBR.

detection. These studies suggest that the blue/red ratio is also important for cloud cancellation and halo extraction in an image.

3 Existing algorithms

This section describes the existing algorithms for extracting halos in images, i.e., $B - R$ processing and autoBR. It also shows how color pixels are translated into grayscale pixels in translation maps.

3.1 $B - R$ processing

$B - R$ processing, also referred as “color subtraction,” is a widely used technique for enhancing and explaining halos and other atmospheric optical phenomena\(^2\). It is introduced as an image processing technique using a tool called the “channel mixer.” It is available in Photoshop or other such photo-retouching applications. The channel mixer outputs a new pixel value of a specified channel calculated from the RGB values of the corresponding pixel in a source image. The values are added or subtracted with specified weights, and a “shift” value is used to adjusts the output range to keep the pixel values in the interested area of interest within 0–255.

For $B - R$ processing, the weight for a red channel is set to the same magnitude but with opposite sign as the weight for a blue channel. Although the magnitude is adjustable, it is often set to 200%, the maximum value allowed by the tools. The shift value is set and adjusted by the user by looking at the resulting image. Figure 3a shows a resulting gray map when the parameters are set to $B$: 200%, $G$: 0%, $R$: $-200\%$, and shift: 0. The blue and red lines show the edge of the underexposed and overexposed area, respectively, while the magenta line shows pixels in the center value (128). The map depicted in Fig. 3b shows the case when the shift parameter is set to $-128$.

The pixel-value translation in $B - R$ processing is represented as follows:

$$I = \alpha (B - R) + \beta, \quad (1)$$

where $I$ is the output intensity value, $B$ and $R$ are pixel values in the source image, and $\alpha$ and $\beta$ are the “magnitude” and “shift” ($I$ is set to 0 if $I < 0$ and to 255 if $I > 255$). For larger $\alpha$, the contrast of the output image will be higher and the width of the area between the blue and red lines in a gray map will be narrower.

3.2 autoBR

Extending $B - R$ processing, more effective parameters have been heuristically explored and the automatic adjustment of the “shift” parameter has been implemented for the autoBR algorithm. The pixel-value translation in autoBR is represented as follows:

$$I = \alpha (\omega_B B + \omega_G G - R) + \beta, \quad (2)$$

where $\alpha = 2.0$, $\omega_B = 1.5$, and $\omega_G = 0.25$. The $\beta$ is set to a value such that the average of the $I$ values (excluding too dark/bright areas in the source image) becomes 128. An example gray map for the translation is shown in Fig. 3c, where the blue, red, and magenta lines are steeper than those for $B - R$ processing.

4 Model and validation

Considering how the presented algorithms work effectively to extract halos, a model of sky color in images can be assumed. This section explains the model and validates it with some photographs.

4.1 A model for sky color

This study provides a simple physical model for celestial light, with no direct beam from the light source and no particles other than those that form clouds. Light $L_{\text{sky}}$ from one direction in the sky contains the following (Fig. 4):
1. Rayleigh scattering $L_R$ from air molecules (sky blue background)

2. Mie scattering $L_M$ from cloud particles (white cloud)

3. refracted/reflected light $L_H$ from cloud particles (e.g., halos).

In a digital image, $L_{sky}$, $L_R$, $L_M$, and $L_H$ are vectors typically containing $R$, $G$, and $B$ intensities. If translation $g$ that converts colors in an image to grayscale reduces the contrast of the cloud color, it satisfies the following formula:

$$g(L_R + L_M + L_H) \approx g(L_R + L_H).$$  (3)

Considering the color-to-gray translation maps in Fig. 3, the above formula implies that the direction of vector $L_M$ is almost parallel to the red, blue, and magenta lines when projected onto the $B$ vs. $R$ space. Figure 5 shows the model for sky color in the $B$ vs. $R$ space. $L_R$ in Fig. 5 represents sky color if the clouds and halos are not there. The arrows represent vectors, $L_M$ and $L_H$, added to the vector $L_R$ (in the $B$ vs. $R$ space). The color of the sky at a given point in an image is the result of the summation of three vectors.

### 4.2 Validation of the model

Figure 6 shows a halo display image and its color maps of it in $B$ vs. $R$, $R$ vs. $G$, and $G$ vs. $B$ spaces. The pixel colors in the maps other than black represent (one of) the real colors in the image. The colors are distributed linearly in the maps, roughly parallel to the lines shown in Fig. 3. Some colors are outside the main distribution, corresponding to the silhouette of the foreground objects in the image. The width of the main distribution is as narrow as the area between the blue and red lines depicted in Fig. 3. These reasons explain the effectiveness of $B - R$ processing and autoBR.

The lines in the color maps are regression lines for the distributions of color in each space. From the slope of the $B$ vs. $R$ regression line, it can be expected that autoBR with a higher slope would be more effective than the $B - R$ method with a slope of 1.0. Blue minus green is a variation in $B - R$ processing (Juutilainen, 2015) and will be effective when the slope of the regression line in the $B$ vs. $G$ space is close to 1.0.

Figures 7 and 8 show partial color maps and their regression lines for the areas indicated by box lines. The partial color maps, corresponding to the red and yellow boxes, are superimposed on the full color maps shown in dimmed colors.

Figure 7 shows the color maps and regression lines for areas without halos. These maps are thinner and more linear than the full map. This means that the blue of the sky, whiteness of the light, and shadow of the clouds are also uniform in these small areas. The slopes of the regression lines for the red and yellow boxes are clearly different from each other,
Such features are less significant in the G color map, indicating the red color of the circumzenithal arc.

Let $\omega$ be adjusted locally rather than using one set of parameters for all areas of an image. Based on these observations, a new algorithm called “sky-color regression” is constructed. For this study, the following conditions were set for the new algorithm: $\alpha_1 = 2.5$ and $\alpha_2 = 1.0$. However, it has not yet been discussed how the green channel affects the sky color; therefore, the values were explored heuristically. A value of $\beta$ was set as in autoBR, i.e., the average of $I$ should be 128. This process is called “regression-based processing”.

Figure 9 shows the images processed with autoBR and regression-based processing. When regression-based processing is applied to the entire image, clouds around the circumzenithal arc are reduced more than that achieved with autoBR, although the former method seemed to be less effective for the other areas. When regression-based processing was applied to a part of the image, clouds in the part were effectively reduced, and the brightness was also adjusted accordingly.

## 5 Sky-color regression algorithm

The color maps explained in the previous section suggest that the RGB color to grayscale is considered as follows:

$$I = \alpha_1 (\omega_{BR} B - R) + \alpha_2 \omega_{RG} \omega_{GB} G + \beta.$$  \hspace{1cm} (4)

For best results, regression parameters $\omega_{BR}$, $\omega_{RG}$, and $\omega_{GB}$ should be updated for every pixel in an image; however, this is computationally expensive. If the best parameters for neighboring pixels are similar, some calculations can be omitted as shown below:

1. The parameters are updated per block of a few pixels.
2. The regressions are performed on an appropriately shrunken image.

When implementing the sky-color regression algorithm, a target image should be divided into $b \times b$-px blocks. The regression parameters are calculated on a shrunken image of size $1/b$. The default value of $b$ is set so that the long side of the shrunken image is approximately 1000 px, i.e., $b = \lceil p/1000 \rceil$, where $p$ is the long side of the target image. If the camera angle is known, the value can be optimized using this information. How to calculate the best value for each angle is a future work of this study.

An example of an image processed by the sky-color regression algorithm is shown in Fig. 10. The width and height of the image are 4024 × 6048 px, and the image was processed in 4.3 s using sky-color regression implemented in Java on a PC equipped with an Intel Core i7-1165G7 (2.80 GHz).

More examples are shown in Fig. 11, which compares sky-color regression with $B - R$ processing and autoBR. While the algorithms mainly enhance colored halos, a colorless formation from the RGB color to grayscale is considered as follows:

$$I = \alpha_1 (\omega_{BR} B - R) + \alpha_2 \omega_{RG} \omega_{GB} G + \beta.$$  \hspace{1cm} (4)

For best results, regression parameters $\omega_{BR}$, $\omega_{RG}$, and $\omega_{GB}$ should be updated for every pixel in an image; however, this is computationally expensive. If the best parameters for neighboring pixels are similar, some calculations can be omitted as shown below:

1. The parameters are updated per block of a few pixels.
2. The regressions are performed on an appropriately shrunken image.

## 5.2 Local adaptive processing

If the area to calculate the regression line is too small, the accuracy of the line will worsen. However, if the area is too large, the locality will deteriorate. Assuming a wide-angle image such as the photos shown in this paper, the default value was set to one-sixth of the long side of an image.

For best results, regression parameters $\omega_{BR}$, $\omega_{RG}$, and $\omega_{GB}$ should be updated for every pixel in an image; however, this is computationally expensive. If the best parameters for neighboring pixels are similar, some calculations can be omitted as shown below:

1. The parameters are updated per block of a few pixels.
2. The regressions are performed on an appropriately shrunken image.

When implementing the sky-color regression algorithm, a target image should be divided into $b \times b$-px blocks. The regression parameters are calculated on a shrunken image of size $1/b$. The default value of $b$ is set so that the long side of the shrunken image is approximately 1000 px, i.e., $b = \lceil p/1000 \rceil$, where $p$ is the long side of the target image. If the camera angle is known, the value can be optimized using this information. How to calculate the best value for each angle is a future work of this study.

An example of an image processed by the sky-color regression algorithm is shown in Fig. 10. The width and height of the image are 4024 × 6048 px, and the image was processed in 4.3 s using sky-color regression implemented in Java on a PC equipped with an Intel Core i7-1165G7 (2.80 GHz).

More examples are shown in Fig. 11, which compares sky-color regression with $B - R$ processing and autoBR. While the algorithms mainly enhance colored halos, a colorless formation from the RGB color to grayscale is considered as follows:

$$I = \alpha_1 (\omega_{BR} B - R) + \alpha_2 \omega_{RG} \omega_{GB} G + \beta.$$  \hspace{1cm} (4)

For best results, regression parameters $\omega_{BR}$, $\omega_{RG}$, and $\omega_{GB}$ should be updated for every pixel in an image; however, this is computationally expensive. If the best parameters for neighboring pixels are similar, some calculations can be omitted as shown below:

1. The parameters are updated per block of a few pixels.
2. The regressions are performed on an appropriately shrunken image.
Figure 6. Color maps of a sky image including halos.

Figure 7. Color maps of areas without halos in the sky image: color maps for the red box are shown in the left column, and those for the yellow box are shown in the right column.
Figure 8. Color maps of areas with halos in the sky image: color maps for the red box, including a parhelion, are shown in the left column, and those for the yellow box, including a circumzenithal arc, are shown in the right column.

Figure 9. Comparison of regression-based processing with autoBR.
Discussion

6.1 Evaluation

Figure 12 displays histograms of processed image parts by the sky-color regression, $B-R$ processing, and autoBR. Rows (a) and (c) show areas without halos, while row (b) shows an area with circumzenithal and supralateral arcs.

With the sky-color regression and autoBR, distributions of pixel values for areas (a) and (c), without halos, are simple standard distributions, while there are two or more peaks in distributions with $B-R$ processing. Standard deviations are also smaller with the sky-color regression than with $B-R$ processing. For area (a), the standard deviation is smaller when using the sky-color regression compared to autoBR. Conversely, for area (c), the standard deviation is smaller with autoBR than with the sky-color regression. It shows that the sky-color regression is not always the most effective method for canceling out cloud and should be refined in future studies.

For area (b), all algorithms produced a histogram with three peaks corresponding to the clouds, reddish parts, and bluish parts of the halos. However, the peaks are most clearly separated with the sky-color regression. The standard deviation is also the largest with the sky-color regression.

The average values for the areas are maintained around 128, which is the midpoint value of the range 0 to 255, with the sky-color regression. It shows that the local adaptive processing works.

6.2 Other atmospheric optical phenomena

The sky-color regression algorithm can also be used efficiently to enhance other colored atmospheric optical phenomena. Figure 13a shows an example of quite faint third- and fourth-order rainbows. The algorithm extracts colored bows from the background sky without clouds. Figure 13b shows the diffraction pattern of Japanese cedar pollen corona. The algorithm also works on the sky color near sunset.

6.3 Limitations

In the sky-color regression algorithm, the mixing of the red and blue channels is calculated statistically, while the weight of the green channel is adjusted heuristically. The size of the pixel block and the area over which the regression is performed are also determined by heuristics. These parameters would also be statistically determined if further analysis was performed on the images of halo displays.

The color maps of an image change according to the white balance setting. Because of its adaptive nature, sky-color regression is less sensitive to white balance than $B-R$ processing and autoBR are. Because digital cameras often have more color settings than white balance, how the settings affect the results of the algorithms needs to be examined.

The sky-color regression algorithm has been developed and optimized mainly for the images captured by standard digital single-lens reflex (DSLR) cameras equipping lenses with focal lengths of 24 to 100 mm. Adjusting parameters may be necessary for other types of images, such as total sky images. Optimizing and evaluating the algorithm for total sky images is an area for further study.
Figure 11. Examples of sky-color regression in comparison to $B - R$ processing and autoBR.
Figure 12. Histograms of processed image parts: rows (a) and (c) are areas without halos, while row (b) is an area with circumzenithal and supralateral arcs.

Figure 13. Applying sky-color regression for images of other atmospheric optical phenomena: (a) third- and fourth-order rainbows and (b) pollen corona (Japanese cedar).

7 Conclusions

This paper considered an image processing method to extract atmospheric optical phenomena such as halos by reducing the contrast of clouds according to their color in the sky image and constructed a model of sky color. It also validated the model with halo images and implemented a new algorithm named sky-color regression. The algorithm extracts halos, including those that are difficult to distinguish with the naked eye. Even the most popular 22° halos are sometimes camouflaged by clouds. In such cases, the new algorithm might allow automated observations. In processed images, halos, which are often geometrically shaped, are more easily detected by feature-detection algorithms. The algorithm is use-
ful not only for accurately observing the appearance of halos, but also to measure the intensity of halos by removing cloud contrasts.

Although the purpose of this work was to process sky images for easier halo detection, the model constructed and validated in this paper and the method developed can be used for other purposes as well as cloud detection. Similar models and processing can be developed that would benefit other fields.


**Data availability.** All the photographs used in this manuscript can be accessed at https://doi.org/10.5281/zenodo.12181277 (Ayatsuka, 2024).

**Competing interests.** The author has declared that there are no competing interests.

**Disclaimer.** Publisher’s note: Copernicus Publications remains neutral with regard to jurisdictional claims made in the text, published maps, institutional affiliations, or any other geographical representation in this paper. While Copernicus Publications makes every effort to include appropriate place names, the final responsibility lies with the authors.

**Acknowledgements.** I thank RainbowMustache (@HIGEHIG-Erainbow on Twitter) for introducing me to $B-R$ processing and Kentaro Araki for advising me on where to submit this paper.

**Review statement.** This paper was edited by Alyn Lambert and reviewed by four anonymous referees.

**References**


