Dehazing of Satellite Images

Johanna Hultberg
Master of Science Thesis in Electrical Engineering

Dehazing of Satellite Images
Johanna Hultberg
LiTH-ISY-EX--18/5121--SE

Supervisor: Hannes Ovrén
ISY, Linköpings universitet
Johan Bejeryd
Vricon Systems AB

Examiner: Lasse Alfredsson
ISY, Linköpings universitet

Computer Vision Laboratory
Department of Electrical Engineering
Linköping University
SE-581 83 Linköping, Sweden

Copyright © 2018 Johanna Hultberg
Abstract

The aim of this work is to find a method for removing haze from satellite imagery. This is done by taking two algorithms developed for images taken from the surface of the earth and adapting them for satellite images. The two algorithms are Single Image Haze Removal Using Dark Channel Prior by He et al. and Color Image Dehazing Using the Near-Infrared by Schaul et al. Both algorithms, altered to fit satellite images, plus the combination are applied on four sets of satellite images. The results are compared with each other and the unaltered images. The evaluation is both qualitative, i.e. looking at the images, and quantitative using three properties: colorfulness, contrast and saturated pixels. Both the qualitative and the quantitative evaluation determined that using only the altered version of Dark Channel Prior gives the result with the least amount of haze and whose colors look most like reality.
Acknowledgments

Without the help from Vricon Systems AB this work would not have been possible. They made me feel welcome from day one and everyone has been very helpful with any problems I have encountered during the process. I would like to give a special thanks to my supervisor, Johan Bejeryd, who put up with all of my questions and helped me no matter the size of the problem.

I would also like to thank my supervisor at the university, Hannes Ovrén, without whom this work and especially the report would have been a lot harder to get done.

Linköping, April 2018
Johanna Hultberg
Contents

1 Introduction ................................................................. 1
  1.1 Goals ........................................................................... 3
  1.2 Data ............................................................................ 3
  1.3 Notation ....................................................................... 3
  1.4 Delimitations ............................................................... 4

2 Theory ............................................................................... 7
  2.1 Atmospheric Correction .................................................. 8
  2.2 Dehazing ....................................................................... 9
  2.3 Methods ....................................................................... 10
    2.3.1 FLAASH .............................................................. 11
    2.3.2 DG AComp .......................................................... 11
    2.3.3 He et al. ............................................................... 11
    2.3.4 Pan et al. .............................................................. 12
    2.3.5 Schaul et al. .......................................................... 14

3 Method ............................................................................... 19
  3.1 Dark Channel ............................................................... 19
  3.2 NIR ............................................................................... 24
  3.3 Combination .................................................................. 25

4 Results and Evaluation ...................................................... 27
  4.1 Evaluation Methods ...................................................... 29
    4.1.1 Qualitative Evaluation Method .................................. 29
    4.1.2 Saturated Pixels ...................................................... 30
    4.1.3 Perceived Colorfulness ............................................. 30
    4.1.4 Contrast Gain ........................................................ 30
  4.2 Dark Channel ............................................................... 31
  4.3 NIR ............................................................................... 33
  4.4 Combination ................................................................. 35
  4.5 Qualitative evaluation ................................................... 37
  4.6 Quantitative evaluation .................................................. 37
5 Discussion
  5.1 Results ................................................. 39
  5.2 Methods ................................................. 42
    5.2.1 Dark Channel ..................................... 42
    5.2.2 NIR ................................................. 43
    5.2.3 Evaluation methods ................................. 43
  5.3 Wider Perspective ..................................... 45

6 Conclusions ................................................. 47
  6.1 Future Work ............................................. 48

A Image Comparison ........................................... 51

Bibliography .................................................. 59
In recent years there has been a growing interest in using satellite imagery for different tasks in modelling the earth, e.g. creating accurate maps. For this to be possible the satellite images must have a high enough resolution to be able to detect small details. Today there exist commercial satellites with a resolution down to 0.31m [6]. The usability of the images does not only depend on their resolution but heavily on the quality of their content as well.

One great disadvantage when taking images from outside of the atmosphere is that the result will be distorted since the light is forced to pass through particles in the atmosphere which will scatter and absorb the light. The distortion will reduce the applicability of the resulting images if they are not corrected. Since these effects are inevitable in the field of remote sensing it is important to have effective methods to remove them.

It can be fairly straight forward to remove the homogeneous atmospheric effects in an image, the scenario in Figure 1.1 (a), simply by subtracting the lowest intensity in each color band. This should correspond to translating the image so that the darkest part is black and not the color that the haze changes it into.

If there is additional fog or haze in the scene, as in Figure 1.1 (b), it has to be compensated for and removed for the result to be as close to the truth as possible. The haze will distort the detected colors in the images and removing it should make the colors more vivid and closer to the truth. The loss of colorfulness in the image also decreases the contrast in the image since the haze has a smoothing effect on the scene. The wanted result for the dehazing is a clear image, as in Figure 1.1 (c), where there is no haze at all in the image.
Introduction

Much work has previously been done in the area of dehazing, some of which will be discussed in section 2.3. By looking at said previous work two algorithms were selected, investigated closer and improved upon. One algorithm, by He et al. [14], uses the dark channel, the smallest value across the RGB-channels of each pixel, of the images to determine the color and the distribution of the haze. The other, by Schaul et al. [20], uses the near-infrared channel of the image to enhance the contrast in the hazy areas of the image. The chosen methods are discussed in more detail in section 2.3.3 - 2.3.5 and the implementations are described in section 3.

Figure 1.1: Scenes containing different amounts of haze.
1.1 Goals

The aim of this study is to find a way to automatically remove haze from satellite imagery. This will be done by implementing two different algorithms and comparing the results. The implementations and the comparisons should answer the following questions:

- Which of the suggested methods gives the best result and is most repeatable for removing haze and atmospheric effects from satellite images?
- Are there conditions not suitable for the suggested methods?
- Which of the suggested methods gives colors closest to the reality after removing the haze?
- Which are the main challenges with dehazing? E.g.:
  - Constant haze in the scene
  - Opaque clouds
  - Different scenes, e.g. desert

1.2 Data

The images used for evaluation are taken over urban areas, forests, deserts and islands. This is to be able to compare how well the algorithm works on different types of landscape. The images are taken with Digital Globe’s satellites GeoEye 1 [3], QuickBird 2 [4], WorldView 2 [5] and WorldView 3 [6] which have a varying number of image bands. In this study five of those bands are used: the three RGB-bands, the NIR-band and the panchromatic band. The RGB-bands are the bands corresponding to the red, green and blue color channel of the image, the NIR-band holds the Near-Infrared spectrum and the panchromatic band holds a gray scale version of the image. The resolution of the images are approximately 0.5 m for the panchromatic band and 2 m for the RGB- and NIR-bands. [5] The reason for only focusing on the five previously mentioned bands is that the developed method is supposed to work regardless of which satellite the images comes from and these are the most common bands.

1.3 Notation

The mathematical notation in this work will follow the convention in Table 1.1. Multidimensional matrices will be in bold upper case letters, regular matrices will be upper case letters, vectors will be bold lower case letters and scalars will be regular lower case letters.

Exponents will be used in three different ways. In Section 2.3.3 the exponent will be used to distinguish between the three channels in the RGB-image and in
the atmospheric color when creating the dark channel image. It will be used in the same way when calculating the mean atmospheric color in Section 2.3.4. The exponents $a$ and $d$ will be used to distinguish between the approximate image and the image holding details in Section 2.3.5. In the cases with an exponent that are not explicitly mentioned here the exponent has its regular meaning, i.e. "to the power of".

**Table 1.1: Mathematical notation.**

<table>
<thead>
<tr>
<th>Notation</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A$</td>
<td>Multidimensional Matrix</td>
</tr>
<tr>
<td>$A$</td>
<td>Matrix</td>
</tr>
<tr>
<td>$\mathbf{a}$</td>
<td>Vector</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>Scalar</td>
</tr>
</tbody>
</table>

### 1.4 Delimitations

The resulting method is not supposed to see through or remove clouds and thick haze, through which the ground can not be seen at all. It should be able to compensate for haze where the ground is visible through the haze, as is marked in Figure 1.2. This is since in the regions with thick haze there is not enough information about the ground below to be able to obtain a satisfying result. The satellite orbits are synchronous with the sun so all of the available images are taken at approximately the same time of the day. Therefore this work will only be tested on images with similar light and shadows. The method is assumed to work with single satellite images and not use any overlapping parts of images for better estimations.
1.4 Delimitations

Figure 1.2: Satellite image containing both haze and opaque clouds. The red circle shows the type of haze studied in this work.
The received image in a satellite can be modeled as

\[ I(x) = J(x)T(x) + (1 - T(x))a. \]  \hspace{1cm} (2.1)

\( I \) and \( J \) are RGB-images with three channels, \( T \) is a matrix with the same x- and y-dimension as \( I \) and \( J \) and \( a \) is a 3x1 vector. Here \( I \) is the observed intensity, \( J \) is the scene radiance and \( a \) is the atmospheric light, which is the same for all pixels in the same band [14]. \( T \) is the transmission of light that is not scattered and reaches the sensor and it is the same for all the bands of the image and \( x \) is a pixel. This equation is calculated pixel-wise, where \( I(x) \), \( J(x) \) and \( a \) are 3x1 vectors and \( T(x) \) is a scalar in each computation. The goal is to extract \( J \) from this equation, which is the image of the canopy without the effects from the atmosphere.

Light of different wavelengths are absorbed and scattered differently in the atmosphere. Shorter wavelengths are scattered more than longer wavelengths and apart from the wavelength it is mostly the amount of \( O_2 \) and \( N_2 \) that affects the light. Differences in the atmosphere conditions also play a part in the atmospheric effects, e.g. the amount of aerosols, water vapor and clouds as well as the angle with which the satellite takes the images will affect the resulting images. The more particles there is between the satellite and the ground the smaller the transmission, i.e. the value of \( T \) is smaller. An angle which gives a longer path for the rays or if there is much aerosols in the atmosphere will result in an image \( I \) that looks hazier than one taken with an angle that gives a shorter path or if there is less aerosols in the atmosphere. The consequence of varying angles will be described further in section 2.1.
2.1 Atmospheric Correction

The resulting intensity $I(x)$ for a pixel in an image of the earth, taken from a satellite, can be described using three different components. First there is the light that has been reflected from the area of interest i.e. the reflected radiation $J(x)$. Secondly there is sunlight that has been reflected in the atmosphere and never touches the surface of the earth i.e. the path radiance. Lastly there is the radiation that comes from the adjacent areas. In Figure 2.1 the different components and how they are constructed can be seen. In this figure $L_{PATH}$ is the path radiance, $L_{WR}$ is the wanted reflected radiance and $L_{NR}$ is the reflected radiance from the neighborhood.

![Figure 2.1: How a remotely sensed image is composed. $L_{PATH}$ is the path radiance, $L_{WR}$ is the wanted reflected radiance and $L_{NR}$ is reflected radiance from the neighbourhood. The black part is the wanted area.](image)

The sunlight that reaches the satellite after it has been reflected by the earth has to go through the atmosphere twice. That will result in both path radiance and contributions from neighbouring areas. It is only the direct reflection that is wanted so in order to get the true image the other contributors must be removed or compensated for. Atmospheric correction is the task of removing the path radiance and the radiance reflected from the neighbourhood to extract only the radiation reflected from the pixel of interest [19]. In (2.1) $J$ corresponds to $L_{WR}$ and $T$ corresponds to the atmosphere. The two other contributors, $L_{NR}$ and $L_{PATH}$, can be compensated for with the second part of (2.1). There, $a$ is the color added from $L_{NR}$ and $L_{PATH}$ and $T(x)$ is the amount of $a$ added to each pixel. The same color is added to all pixels in the same image, it is the amount that is added that differs between pixels.

Rayleigh scattering occurs when light collides with particles that are much smaller than its wavelength, compared to larger particles which give rise to Mie
scattering. The intensity of scattered light is inversely proportional to the fourth power of the wavelength, which means that blue light is affected the most since it has the shortest wavelength. The amount of Rayleigh scattering also depends on the observation angle from satellite to ground. The greater the angle, the more atmosphere the rays have to go through and the more particles they have to pass. In Figure 2.2 the difference in path for two observation angles can be seen. The longer path in the left figure results in more possibilities for the atmosphere to scatter and absorb the light. Rayleigh scattering depends on the amount of particles, e.g. nitrogen and oxygen molecules, in the air. Since a higher pressure in the atmosphere means more particles to possibly collide with, the atmospheric pressure affects the formation of the images. The thickness of the ozone layer also affects the imagery since a thicker layer means more particles. [19] All these factors result in a condition where the amount of atmosphere to be removed or compensated for varies between all images. Because of this, there is no set amount to remove from every image, so the computations have to be done for each image individually. Even between images taken of the same area at approximately the same time there are some small differences, so for the best result there has to be separate computations for each image.

Figure 2.2: The difference in path length for different observation angles from satellite to ground, the angles are marked in both figures.

2.2 Dehazing

Haze comes from particles in the air that scatters and obscures the light. It consists of e.g. dust, smoke, water vapor or other small particles in the atmosphere. Dehazing is the process of removing thicker layers of haze not created only from the light scattering in the atmosphere, but from e.g. water vapor. The haze is nonuniform in comparison to the atmospheric scattering. In some areas the ground is visible, but partially obscured by a layer of haze, as can be seen in the marked area in Figure 1.2. This is the form of haze that this work aims to compensate for.

The thickness of the haze and how it is distributed is estimated using the
transmission map $T$. By removing different amounts, depending of the value of $T(x)$, of the estimated atmosphere $a$ from the initial image $I$, an estimate of the true scene $J$ can be received.

If there is no additional haze in the image, the atmospheric effect is approximately uniform in the entire image. However, the effect will, as was previously mentioned, vary between images depending on the atmospheric pressure etc, so the amount of haze that is removed will still not be the same for every image.

## 2.3 Methods

Recently there has been much work done in the area of improving satellite images. In the area of dehazing several different approaches has been published [1], [14], [16], [17], [18]. One well known and often cited method is Dark Channel Prior by He et al., which however uses regular images instead of remotely sensed images [14]. "Regular" images here refer to images taken on the surface of the earth, inside of the atmosphere.

One of the simplest methods of removing atmospheric effects is finding the darkest pixel in the image and assuming that it corresponds to black, Dark Object Subtraction. To get the "true" image from this, the value of the "black" pixel is subtracted from each pixel in the image. This corresponds to removing the additive term in (2.1), i.e. it removes the light scattered in the atmosphere that is added. However, this removes spatially homogeneous haze and does not take variations in the haze across the image into consideration. [2]

Dark Channel Prior [14] also uses (2.1). It uses the dark channel, which is the lowest value across all three RGB-channels in each pixel, to calculate the atmospheric light $a$. $a$ combined with the observed image is used to calculate the transmission, $T$, through the haze. This method is more thoroughly described in section 2.3.3.

From the Dark Channel Prior method other methods are derived. In [18] the dark channel of the image is assumed to have an average intensity that is not zero, as assumed in [14], but an average intensity that is translated to a brighter average. This is due to the atmosphere scattering the light and adding the intensity of the haze to the intensity of the scene. This increases the average intensity of the image and by translating the resulting histogram of the dark channel so its average intensity is placed close to zero the assumptions from He et al. [14] can be used. This method is more thoroughly described in section 2.3.4.

One different method that uses the NIR-spectrum and relies on the fact that rays with a wavelength in the NIR-spectrum scatters less than rays in the visible spectrum due to the longer wavelength is suggested by Schaul et al. [20] and is described further in section 2.3.5.
Some complete methods for correction of satellite imagery exist, none of which are optimal. They either require some level of user input or have many parameters to set internally. Two of these methods will be described further in section 2.3.1 and 2.3.2. The methods this work focuses on are described in section 2.3.3, 2.3.4 and 2.3.5.

### 2.3.1 FLAASH

Fast Line-of-sight Atmospheric Analysis of Spectral Hypercubes, FLAASH, is a tool for removing the atmospheric effects, including haze, in remotely sensed images [21]. FLAASH is physics based and is therefore considered to be more accurate than non-physical methods. This means that FLAASH might give better results than other methods but it is also more computational heavy. The biggest downside with FLAASH is that it requires input of the atmospheric conditions from the user, which might not be known for every case.

### 2.3.2 DG AComp

DigitalGlobe Atmospheric Compensation, DG AComp, is a method for correcting satellite imagery developed by DigitalGlobe [10]. It uses a physics based model to correct the images from the effects caused by the atmosphere, not just haze. It produces an Aerosol Optical Depth (AOD) map and compensates for the atmospheric effects on a pixel level. It does not need any input from the user but the model it uses have several internal parameters to set. Therefore the processing time is longer compared to other methods, e.g. FLAASH, since DG AComp itself estimates the parameters.

### 2.3.3 He et al.

Dark Channel Prior [14] uses (2.1) to retrieve the true scene. The dark channel is a monochrome image composed of the value of the channel with the lowest intensity across the RGB-bands in each pixel

\[ I_{\text{Dark}}(x) = \min_{c \in r,g,b} I^c(x). \]  

(2.2)

This can be seen as a distance map over the scene. It can be assumed that the distances in a neighboring region, \( \Omega(x) \), are similar so the map is min-filtered in smaller regions. The dark channel can be translated into a rough transmission map, \( \tilde{T}(x) \), with

\[ \tilde{T}(x) = 1 - \omega( \min_{y \in \Omega(x)} I_D(y)), \]  

(2.3)

where \( I_D \) is a modified version of (2.2) where each channel is divided by the corresponding atmospheric color in \( a \):
\[ I_D(x) = \min_{c \in \{r,g,b\}} \frac{I_c(x)}{a^c}. \] (2.4)

Here the transmission is calculated from the hazy image \( I \) relative to the atmospheric light in each RGB-channel. The constant \( \omega \) is introduced and fixed to 0.95 to keep a small amount of haze in the distant regions to mimic the way humans perceive depth. The transmission is assumed to be constant in a local patch and the smallest value in the patch is wanted, hence it is calculated from the min-value in a neighborhood that is 15x15 pixels, \( \Omega(x) \). The rough transmission from this patch is \( \tilde{T}(x) \).

The atmospheric light, \( a \), for each channel is estimated as the mean value, from the corresponding channel in the input image, of the 0.1% brightest pixels in the dark channel. This method uses the assumption that haze free images have a low intensity in at least one of the RGB channels to create the dark channel.

The rough transmission map, \( \tilde{T} \), is refined using a soft matting algorithm. The refined transmission, \( T \), is computed using

\[ (L + \lambda U)T = \lambda \tilde{T}. \] (2.5)

\( L \) is a Matting Laplacian matrix [15], \( U \) is an identity matrix the size of \( L \) and \( \lambda \) is a small scalar value so \( T \) is softly constricted by \( \tilde{T} \).

The resulting image without haze is calculated as

\[ J(x) = \frac{I(x) - a}{\max(T(x), t_0)} + a. \] (2.6)

Here the denominator uses \( t_0 \), which is fixed to 0.1, to prevent distortions in the resulting image due to very small values in \( T \).

However, this method is created for regular scenes and not for satellite images which results in it not working perfectly for the purpose in this work.

### 2.3.4 Pan et al.

The method suggested by Pan et al. [18] is heavily based on the previously discussed method by He et al. It uses the same basic assumptions of the problem but they use remotely sensed images, instead of the regular images used by He et al. This makes the assumption that the average intensity of the pixels in the dark channel of an image is close to zero not completely true. This problem is solved by introducing a translating term, \(-cT(x)\), to both sides of (2.1) which after some rearranging results in

\[ I(x) - cT(x) = (J(x) - c)T(x) + (1 - T(x))a. \] (2.7)
2.3 Methods

c corresponds to the translation in average intensity in the dark channel between regular images and remotely sensed images. In [18] c is estimated by manually selecting 5000 haze free remotely sensed images from Google Earth and calculating the average intensity of all 5000 dark channel images.

Pan et al. uses a guided filter [13] to refine the rough transmission map, $\tilde{T}$, instead of the Soft Matting Laplacian used in [14]. The guided filter uses the hazy input image, $I$, as a guide to preserve the details of the scene when filtering the transmission map. The filter uses the model

$$\tilde{T}(x) = b(x)^T I(x) + d(x). \quad (2.8)$$

Here $\tilde{T}(x)$ is a first estimate of the refined transmission map, $I(x)$ is a 3x1 vector of pixel $x$ in the guiding image and $b(x)$ and $d(x)$ are as follows:

$$b(x) = (\Sigma(x) + \epsilon U)^{-1} \left( \frac{1}{|\Omega(x)|} \sum_{y \in \Omega(x)} (I(y) \tilde{T}(y) - \mu(x) \tilde{T}(x)) \right), \quad (2.9)$$

$$d(x) = \overline{\tilde{T}}(x) - b(x)^T \mu(x). \quad (2.10)$$

$\tilde{T}(x)$ is the value of the rough transmission map in pixel $x$, $\mu(x)$ and $\Sigma(x)$ corresponds to the 3x1 mean vector and the 3x3 covariance matrix of $I$ in the neighborhood $\Omega(x)$. $|\Omega(x)|$ is the number of pixels in $\Omega(x)$ and $U$ is a 3x3 identity matrix. $\overline{\tilde{T}}(x)$ is the mean of $\tilde{T}$ in $\Omega(x),$

$$\overline{\tilde{T}}(x) = \frac{1}{|\Omega(x)|} \sum_{y \in \Omega(x)} \tilde{T}(y). \quad (2.11)$$

To reduce noise in Equation 2.8, the coefficients $b$ and $d$ are averaged from all windows overlapping pixel $x$:

$$\bar{b}(x) = \frac{1}{|\Omega(x)|} \sum_{y \in \Omega(x)} b(y) \quad (2.12)$$

and

$$\bar{d}(x) = \frac{1}{|\Omega(x)|} \sum_{y \in \Omega(x)} d(y). \quad (2.13)$$

$\overline{\tilde{T}}(x)$ and $\bar{d}(x)$ are scalars and $\bar{b}(x)$ is a 3x1 vector containing the coefficients for each channel of the image. This results in the filtered and refined transmission map

$$T(x) = \bar{b}(x)^T I(x) + \bar{d}(x). \quad (2.14)$$

To retrieve the true image (2.6) is used in this method as well. However, Pan et al. assumes the haze to be gray and that the atmospheric light therefore can be
considered equal in all three channels. This results in the atmospheric light $a$ for each channel being fixed as the mean of the channels $a_0$ and $a$ is

$$a = a_0 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix},$$

(2.15)

$$a_0 = \frac{1}{3} \sum_{k \in r,g,b} a^k.$$  

(2.16)

### 2.3.5 Schaul et al.

The two previously mentioned methods only use data from the visible spectrum but there is useful information in the near-infrared, NIR, spectrum that they miss. Since the rays in the NIR-spectrum have longer wavelengths than the rays in the visible spectrum they don’t scatter as much. Therefore the image in the NIR-spectrum holds more detail than the visible band in hazy regions [20], this can clearly be seen in Figure 2.3. Since many of the commercial satellites today capture images in the NIR-spectrum as well as in the visible spectrum the data for this method is easy to acquire.

The property that there is more detail in a NIR-image in hazy regions is used by Schaul et al. [20] in order to remove haze in images. Their approach is to modify the luminosity channel, the Y-channel, in the YUV-color space with the help of the intensity image in the NIR-channel. This should increase the contrast in the hazy areas and leave the clear areas untouched. This is done without the need to construct a depth map or approximate the atmospheric light as in the previous methods.

The luminance channel of the input image, $V_0$, and the NIR intensity image, $N_0$, are decomposed into contrast pyramids where the index $k$ is the level in the pyramid an image corresponds to. The images are decomposed into images holding the details $I^a_k$, in this case the contrast, and the approximate image, $I^a_k$, separately as shown in Figure 2.4. The images used for the resulting representation are the ones in the blue field. Each new level in the pyramid is computed using

$$I^a_{k+1} = W_{\lambda_0 e^k}(I_0)$$

where

$$W_A(G) = (U + \lambda L_G)^{-1} G,$$

(2.17)

where $\lambda = \lambda_0 e^k$. 

(2.18)

In this expression $U$ is an 3x3 identity matrix and $L_G$ is a 3x3 matrix that is re-calculated for each pixel in $G$. It is then applied to a 3x3 section of $G$ surrounding the current pixel. $L_G$ is computed using

$$L_G = D_x^T A_x D_x + D_y^T A_y D_y.$$ 

(2.19)
2.3 Methods

(a) Blue channel

(b) NIR-channel

Figure 2.3: The blue (a) and NIR (b) channels of the same image.

Here $D_x$ and $D_y$ are difference operators and $A_x$ and $A_y$ are diagonal matrices holding smoothness weights depending on the input image $G$ [9]. The difference operators, $D_x$ and $D_y$, from (2.19) helps choose the proper values from the derivatives. The applied form of (2.19) is therefore

\[
L_G(x, y) = \begin{bmatrix}
0 & A_y(x, y - 1) & 0 \\
A_x(x - 1, y) & 2(A_x(x, y) + A_y(x, y)) & A_x(x + 1, y) \\
0 & A_y(x, y + 1) & 0
\end{bmatrix}.
\] (2.20)

$A_x$ and $A_y$ hold weights for the smoothing of the different layers and they are
calculated by modifying the X- and Y-derivative of the logarithm of the channels, respectively. The $A_x$ matrix holding the modified X-derivative for $V_0$ is calculated with

$$A_x(x, y) = \left( \left| \frac{\delta \log(V_0)}{\delta x}(x, y) \right| + \epsilon \right)^{-1},$$  \hspace{1cm} (2.21)$$

the corresponding matrix $A_y$ and the equivalent matrices for $N_0$ are calculated in the same way. Here $\epsilon$ is a small constant to prevent division by zero in constant areas of the image and it is set to 0.1. The derivatives in (2.21) are calculated using 3x3 Sobel filters.

(2.18) is an edge preserving filter from which the result will become progressively coarser with the increase of $\lambda$. $\lambda = \lambda_0 c^k$ determines the coarseness of the approximate image at level $k+1$, $I_{a_{k+1}}$, and $c$ is a constant. The image holding the contrast is calculated pixel-wise from the approximate images from the current and previous level using

$$I_k^d = \frac{I_{a_{k-1}} - I_k^a}{I_k^a}.$$  \hspace{1cm} (2.22)
The idea is to replace the prior luminance with one calculated from

\[ V_{\text{new}} = V_n^d \prod_{k=1}^n \left( \max(V_k^d, N_k^d) + 1 \right). \] (2.23)

This relies on the assumption that there is more contrast in the hazy regions in the NIR image than in the visible. To achieve the resulting image with the highest contrast possible the two contrast images, \( V_k^d \) and \( N_k^d \) for the visible luminosity respective the NIR-channel, are combined in a way that pixel-wise chooses the maximum value of the two. This should leave the input image unaltered in the clear regions and improved in the hazy regions. This means that this method can be applied on an entire image without changing the clear parts of the image.
Two different methods for dehazing are implemented and compared in this work. The first one is a modified combination of the two methods by He et al. and Pan et al. from section 2.3.3 and 2.3.4 and the second one is based on the method by Schaul et al. in section 2.3.5. These altered methods are described in detail in section 3.1 and 3.2. The methods are implemented in C and the parts not suitable for C, e.g. matrices, are implemented in C++ using the library Eigen [11]. For the guided filter introduced in section 2.3.4 and the Sobel filter in section 3.2 the implementations in OpenCV 3.0 are used [7], [8].

### 3.1 Dark Channel

The assumption that the average intensity of the dark channel is close to zero can be considered true in the settings used in this work. The average intensity can be assumed to be approximately the value that the atmosphere has and in this work it has already been taken care of before the dehazing algorithm is run. This is since the Rayleigh scattering has already been removed in the images used. The dehazing is performed as the last step in a series of steps to develop the images from raw data from the satellites into the images used by Vricon Systems. The removal of the Rayleigh scattering is one of the steps before the dehazing in this process. The input image without the Rayleigh scattering removed can be seen in Figure 3.1 (a) and the same image but with the Rayleigh scattering removed in Figure 3.1 (b). The histograms in Figure 3.2 are of every pixel in the dark channel of the corresponding image in Figure 3.1. From the histograms it can be seen that the mean intensity of the dark channel of the image without the Rayleigh scattering removed is higher than in the image where the Rayleigh scattering is removed. Removing the Rayleigh scattering translates the histogram of the dark
channel from the level of the scattered light down to zero. Therefore (2.1) can be used without the modification for satellite images in section 2.3.4.

Figure 3.1: The input image without (a) and with (b) Rayleigh scattering removed.

The satellite images are larger than the regular images used by He et al. so where they used areas of 15x15 pixels to filter the dark channel it is more appropriate to use 50x50 pixels for satellite images. For the images used in this work 50x50 pixels corresponds to areas of 100x100m. In Figure 3.3 the difference between using 15x15 pixels, (a), and 50x50 pixels, (b), on satellite images can be seen. When using 15x15 pixels much of the details in the scene below the clouds and haze will still be present in the dark channel image. Therefore the coarser filtering with 50x50 pixels will follow the clouds smoother and it will hence result in a more accurate estimation of the haze and cloud cover. The higher the
3.1 Dark Channel

Figure 3.2: Histograms of the intensity of all pixels in the dark channel from images without (a) and with (b) Rayleigh scattering removed.

The intensity of a pixel is the more opaque the haze in that pixel is in the input image. The lightest parts in Figure 3.3 (b) corresponds directly to the parts with cloud in Figure 3.1 (b). The same goes for the lighter gray parts, e.g. bottom left quarter, which corresponds to the thinner haze in the same area in Figure 3.1 (b).

To not use the pixels that are parts of clouds for the determining of the atmospheric light a cloud mask which was developed by Vricon Systems before this work, Figure 3.4, is used. This prevents the estimated color of the atmosphere to be calculated wrong if there is a lot of clouds. Without the cloud mask the 0.1% brightest pixels in the dark channel are chosen, regardless if they correspond to cloud or not. With the cloud mask the 0.1% brightest pixels that have not previously been classified as corresponding to cloud are chosen to determine the atmospheric color from. The difference between using a cloud mask or not can
be seen in Figure 3.5 where the larger white part in (a) directly corresponds to the brightest part in Figure 3.3, which is classified as cloud in Figure 3.4. When comparing Figure 3.5 (b) with Figure 3.4 one can see that all of the white pixels in the mask, located in the bottom right part of the image, are black in the cloud mask. From the pixels masked out in this step the atmospheric light $a$ is calculated. For each of the RGB-channels it is the mean value of the pixels in the input image corresponding to the white pixels in the mask in Figure 3.5 that determines $a$. This gives an RGB-value as the atmospheric light.

The rough transmission map, $\tilde{T}$, is determined from (2.3) which creates the
3.1 Dark Channel

**Figure 3.4:** The cloud mask where white is classified as cloud.

![Cloud Mask](image)

(a) Without cloud mask  
(b) With cloud mask

**Figure 3.5:** The mask used to find which pixels to decide the atmospheric light from.

map in Figure 3.6 (a). This map still holds some detail from the input image so to get a smoother transmission map, $T$, a filter is applied. This is done using a guided filter as done in the method by Pan et al. in section 2.3.4, with a window size of 50x50 pixels and $\epsilon = 0.001$. The result from this is a smoother map, Figure 3.6 (b), which will prevent sharp edges when removing the haze. For the guided filter, the input image is used as the guiding image $I$ and the rough transmission map as the image to be filtered.

From the filtered transmission map, the atmospheric light and the input image the resulting dehazed image is calculated using (2.6). $t_0 = 0.1$ is used, as done by both He et al. and Pan et al, to prevent division by really small numbers.
3.2 NIR

The method suggested by Schaul et al. was implemented as it is described in the paper with one exception. Instead of using the luminosity, which is the Y-channel in the YUV color space, the value-channel in the HSV color space is used for an easier implementation. This can be done since the V-channel and the Y-channel are different representations of the same property, luminosity is a linear combination of the color components of a pixel whereas value is just the largest component.

(2.20) takes the derivatives in the adjacent area to the current pixel into consideration when determining \( L_G \). This is done once for the NIR-channel and once for the V-channel using their respective X- and Y-derivatives since the two channels are significantly different as can clearly be seen in Figure 3.7.
The value for $\lambda_0$ set by Schaul et al. is for regular images and has to be modified for satellite images. A smaller $\lambda_0$ results in a larger difference between the levels since the inverse of $\lambda$ is used in (2.18). So instead of using $\lambda_0 = 0.1$ as suggested in the paper, 0.001 is used. The smaller $\lambda_0$ is useful due to the scale of the satellite images. The larger changes between the levels with the smaller $\lambda_0$ will result in that wanted differences will be detected, ones that could have been missed with the larger $\lambda_0$. The values for $c$ and $n$ are kept as in the paper, as $c = 2$ and $n = 6$.

This method only handles the contrast in the image but it does not compensate for the change in color towards white that the haze contributes to. Therefore this method should be complemented with improving the color in the hazy regions, i.e. removing the white.

### 3.3 Combination

Since the method by Schaul et al. does not handle the discolorations that the haze introduces especially well, a combination of the two methods using dark channel respective NIR-channel is implemented. The result from the algorithm using dark channel, described in section 3.1, is run through the algorithm using the NIR-channel. This is to improve the contrast in the image after the filtering done in the initial algorithm.
Since there is no ground truth for this type of problem the evaluation of the results has to be mainly qualitative instead of quantitative. The main method of evaluation will therefore be to show the resulting images to a group of people and ask them to choose which image they think is the closest to the truth. They will not know which image is the result of which method for their choice to be as unbiased as possible. In section 4.1 this method, and three quantitative properties are described further.

The methods are evaluated on four different types of terrain: desert, urban area, forest and islands. Figure 4.1 shows six images depicting these areas. These images are used to illustrate the effects from the three different algorithms. Figure 4.1 (a), (e) and (f) depict island area where (a) only shows water and no islands, (e) and (f) has island with different amounts of haze. (d) is a good example of a generic image from the desert and (c) is a good example depicting forest with both haze and opaque clouds. The image in (b) is a special example depicting urban area with a lot of homogeneous haze.

First results from the different methods will be shown in section 4.2 - 4.4 and then the evaluation of the results will be presented in section 4.5 and 4.6. In Appendix A the original and the three resulting images are shown side by side for each image for easier comparison between the methods.
Figure 4.1: Images before any of the algorithms.
4.1 Evaluation Methods

Due to the lack of ground truth data the evaluation process has to be mainly qualitative. This is done by using a group consisting of people from Vricon Systems that work with satellite images on a daily basis as an evaluation group. This will result in a subjective evaluation, so to get an objective evaluation as well three properties are compared between the resulting images. For a quantitative evaluation the percentage of saturated pixels, the contrast gain and the perceived colorfulness of each image are compared between the three methods and the original.

4.1.1 Qualitative Evaluation Method

The results from the three different methods and the original image are compared to each other side by side, as is shown in Figure 4.2. Each set is then presented to the evaluation group to determine which of the four images holds the least amount of haze and is the closest to the truth. Four different areas are evaluated, they each represent different conditions as they depict areas in desert, forest, urban area and islands. Even though the goal is to remove as much haze as possible, the image and its colors can not be distorted in the process since that will make the result unusable. Therefore the goal for the evaluation group is to find the image with the colors most like reality and not just the one with the least amount of haze.

Figure 4.2: The setup for the qualitative evaluation. Each image comes from one of the two methods, the combination or is the original.
For each area approximately 15 sets containing four versions of the same image are shown to the evaluation group where they have to choose which of the four is the "best". Where the "best" image is the one with the least amount of haze and whose colors look most like the true colors. Each set of images will have the results from the methods placed in different order so the person evaluating will not know which image corresponds to which method.

4.1.2 Saturated Pixels

The percentage of saturated pixels is measured using

$$\sigma = \frac{n_s}{xy},$$

(4.1)

where $x$ and $y$ corresponds to the width and height of the image and $n_s$ is the number of pixels whose value in the Y-channel of the normalized YUV color space is equal to 0 or 1 [12].

4.1.3 Perceived Colorfulness

The perceived colorfulness of the image is calculated using

$$m = \sqrt{\sigma_\alpha^2 + \sigma_\beta^2 + 0.3\sqrt{\mu_\alpha^2 + \mu_\beta^2}}.$$

(4.2)

Here $\alpha = R - G$ and $\beta = \frac{1}{2}(R + G) - B$, where R, G and B corresponds to respective channel in the RGB color space. $\mu$ is the mean and $\sigma$ is the standard deviation of the $\alpha$- respective $\beta$-channel. [22]

4.1.4 Contrast Gain

The contrast gain is calculated as the difference between the mean contrast in the original image and the dehazed one [23]. The mean contrast for an image is calculated with

$$\bar{C}_I = \frac{1}{xy} \sum_{j=0}^{y-1} \sum_{i=0}^{x-1} \frac{S(i,j)}{E(i,j)}.$$

(4.3)

where $x$ and $y$ are the dimensions of the image and $E(i,j)$ and $S(i,j)$ are calculated using the following two expressions:

$$E(i,j) = \frac{1}{(2p + 1)^2} \sum_{k=-p}^{p} \sum_{l=-p}^{p} I(i+k, j+l).$$

(4.4)
4.2 Dark Channel

\[ S(i, j) = \frac{1}{(2p + 1)^2} \sum_{k=-p}^{p} \sum_{l=-p}^{p} |I(i + k, j + l) - E(i, j)|. \]  

(4.5)

\( p \) is the radius of the window the contrast is calculated within, it is set to 20 in this work which results in a window size of 41x41 pixels. In these expressions I is the Y-channel, in the YUV color space, of the image S and E are calculated for, i.e. the hazy original and the dehazed image, respectively. This results in the contrast metric

\[ c = \bar{c}_{\text{dehazed}} - \bar{c}_{\text{hazy}}. \]  

(4.6)

A large value on \( \bar{c} \) corresponds to a larger contrast in the image. Therefore a larger value on c is wanted since it corresponds to a larger increase in contrast between the original and the result.

These three properties indicates how well the dehazing works. Since haze makes the colors in the image duller the value for colorfulness should be larger for an image containing a smaller amount of haze. Less haze should also increase the contrast in the image so the value of c should be larger after a successful run through the dehazing step. For the result to be good the number of saturated pixels has to be as low as possible, therefore the wanted combination of the properties is a large value for \( m \) and \( c \) and a small value for \( \sigma \).

4.2 Dark Channel

The algorithm using dark channel works well in most areas. For forest where there is both haze and opaque clouds it removes the haze and keeps the rest of the image close to before, only slightly darker as in Figure 4.3 (c). For urban areas with clouds combined with a homogeneous layer of haze covering the entire image the result is also very satisfying, Figure 4.3 (b). In that image the discoloration in the upper left corner comes from removing a lot of haze combined with errors in previous radiometric processing. In areas depicting desert the result is dark, Figure 4.3 (d). This is since the images depicting desert are very sensitive to changes in value after the atmosphere has been removed. The cause for this is that the atmosphere has a large intensity in this area and when it is removed the span of the remaining intensity is very narrow, hence the remainder is sensitive to changes.

For this algorithm to work on areas depicting islands the image must contain some land or something that is not only water. If there is no land the result will be an image that is almost entirely white, Figure 4.3 (a). For the images with islands, (e) and (f), the haze is removed for the most part. The downside is that the residue from the haze makes the resulting image blueish in the parts with thicker haze. This can be seen in the lower right corner in Figure 4.3 (f). The blue color can also be found in the clouds in Figure 4.3 (e).
4 Results and Evaluation

(a) Islands (only water)  
(b) Urban area

(c) Forest  
(d) Desert

(e) Islands  
(f) Islands

**Figure 4.3:** Images after the dark channel algorithm.
This algorithm improves the images depicting desert, it does not remove the haze but it does improve the colors somewhat. This is since there was not a lot of contrast in the original images so even a small improvement will be visible here. In the other areas where there is more contrast to begin with a small change is not as visible. When comparing Figure 4.4 (d) with the corresponding one in Figure 4.1 this one has greater contrast than the unprocessed one. For the images in the other areas, i.e. Figure 4.4 (a) - (c) and (e) - (f), no particular improvement or change can be seen. This is the case for all images in these areas after this algorithm, not just the ones shown here.
Figure 4.4: Images after the NIR algorithm.
4.4 Combination

The combination of the dark channel-algorithm and the NIR-algorithm, Figure 4.5, give results that for most areas are very similar to the ones from only dark channel, Figure 4.3. This is since the NIR-algorithm does not do very much in most areas. This does not apply for the desert area though, Figure 4.5 (d). Here the result is a combination of the results in Figure 4.3 (d) and 4.4 (d) who both makes the result darker. This leads to the combined result becoming much darker than the unprocessed image in Figure 4.1 (d).
Figure 4.5: Images after the combination of the algorithms.
4.5 Qualitative evaluation

The test group consisting of ten persons from Vricon Systems were presented with a series of 61 images as the example in Figure 4.2. The 61 images consisted of 17 island images, 14 desert images, 15 images over urban area and 15 images depicting forest.

Table 4.1 shows the distribution of the choices from the evaluation group of which setting was the best. For the table to have an easy overview, the fields have a darker color the more the corresponding setting was chosen.

**Table 4.1: The result from the qualitative evaluation.**

<table>
<thead>
<tr>
<th></th>
<th>Original</th>
<th>Dark Channel</th>
<th>NIR</th>
<th>Dark Channel + NIR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Desert</td>
<td>13%</td>
<td>64%</td>
<td>19%</td>
<td>4%</td>
</tr>
<tr>
<td>Urban area</td>
<td>34%</td>
<td>24%</td>
<td>20%</td>
<td>22%</td>
</tr>
<tr>
<td>Forest</td>
<td>11%</td>
<td>28%</td>
<td>17%</td>
<td>44%</td>
</tr>
<tr>
<td>Islands</td>
<td>3%</td>
<td>38%</td>
<td>15%</td>
<td>44%</td>
</tr>
</tbody>
</table>

4.6 Quantitative evaluation

For the quantitative evaluation of the results the three properties described in section 4.1 was used. The values presented in Table 4.2 - 4.5 are calculated as the mean value from the images depicting the same area and the standard deviation of the same values. Since the calculation of c uses both the original and the processed image there are no values for c for the originals. In Table 4.2 - 4.5 the best result for each property is marked in bold text.

**Table 4.2: Values for desert.**

<table>
<thead>
<tr>
<th></th>
<th>Original</th>
<th>Dark Channel</th>
<th>NIR</th>
<th>Dark Channel + NIR</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean m</td>
<td>43.802</td>
<td>46.089</td>
<td>41.859</td>
<td>21.187</td>
</tr>
<tr>
<td>mean c</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.001</td>
</tr>
<tr>
<td>std m</td>
<td>6.994</td>
<td>6.944</td>
<td>3.843</td>
<td>5.204</td>
</tr>
<tr>
<td>std c</td>
<td>—</td>
<td>0.007</td>
<td>0.004</td>
<td>0.006</td>
</tr>
</tbody>
</table>
### Table 4.3: Values for urban area.

<table>
<thead>
<tr>
<th></th>
<th>Original</th>
<th>Dark Channel</th>
<th>NIR</th>
<th>Dark Channel + NIR</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean m</td>
<td>16.518</td>
<td>18.676</td>
<td>14.925</td>
<td>18.881</td>
</tr>
<tr>
<td>mean σ</td>
<td>0.007</td>
<td>0.013</td>
<td>0.008</td>
<td>0.012</td>
</tr>
<tr>
<td>mean c</td>
<td>—</td>
<td>0.027</td>
<td>0.003</td>
<td>0.025</td>
</tr>
<tr>
<td>std m</td>
<td>4.002</td>
<td>6.074</td>
<td>3.723</td>
<td>6.121</td>
</tr>
<tr>
<td>std σ</td>
<td>0.007</td>
<td>0.012</td>
<td>0.007</td>
<td>0.011</td>
</tr>
<tr>
<td>std c</td>
<td>—</td>
<td>0.020</td>
<td>0.002</td>
<td>0.020</td>
</tr>
</tbody>
</table>

### Table 4.4: Values for forest.

<table>
<thead>
<tr>
<th></th>
<th>Original</th>
<th>Dark Channel</th>
<th>NIR</th>
<th>Dark Channel + NIR</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean m</td>
<td>15.713</td>
<td>15.618</td>
<td>15.786</td>
<td>15.706</td>
</tr>
<tr>
<td>mean σ</td>
<td>0.016</td>
<td>0.021</td>
<td>0.016</td>
<td>0.020</td>
</tr>
<tr>
<td>mean c</td>
<td>—</td>
<td>0.031</td>
<td>0.001</td>
<td>0.032</td>
</tr>
<tr>
<td>std m</td>
<td>2.910</td>
<td>4.192</td>
<td>3.228</td>
<td>4.480</td>
</tr>
<tr>
<td>std σ</td>
<td>0.009</td>
<td>0.010</td>
<td>0.008</td>
<td>0.010</td>
</tr>
<tr>
<td>std c</td>
<td>—</td>
<td>0.067</td>
<td>0.010</td>
<td>0.074</td>
</tr>
</tbody>
</table>

### Table 4.5: Values for islands.

<table>
<thead>
<tr>
<th></th>
<th>Original</th>
<th>Dark Channel</th>
<th>NIR</th>
<th>Dark Channel + NIR</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean m</td>
<td>18.643</td>
<td>18.853</td>
<td>17.673</td>
<td>18.343</td>
</tr>
<tr>
<td>mean σ</td>
<td>0.025</td>
<td>0.066</td>
<td>0.027</td>
<td>0.064</td>
</tr>
<tr>
<td>mean c</td>
<td>—</td>
<td>0.015</td>
<td>0.009</td>
<td>0.016</td>
</tr>
<tr>
<td>std m</td>
<td>6.514</td>
<td>8.875</td>
<td>6.410</td>
<td>8.495</td>
</tr>
<tr>
<td>std σ</td>
<td>0.037</td>
<td>0.111</td>
<td>0.038</td>
<td>0.110</td>
</tr>
<tr>
<td>std c</td>
<td>—</td>
<td>0.074</td>
<td>0.016</td>
<td>0.074</td>
</tr>
</tbody>
</table>
Most of the results were satisfying in varying degrees. Some images were close to ideal and some were the complete opposite and became unusable. The implemented methods had varying results, where the one using dark channel gave the best results regarding removing the white color of the haze as well as keeping the colors close to reality. This method performed best when considering both the qualitative and quantitative evaluation and just how the images look overall.

5.1 Results

The choices from the evaluation group in Table 4.1 conforms with the observed results, which were briefly mentioned in section 4 and will be discussed more closely below. The values for urban area in Table 4.1 shows that all the settings were chosen approximately the same number of times. This comes from the fact that a substantial part of the images over this area in the evaluation were clear from haze to begin with. For this case none of the methods do any significant change on the image and therefore all four images look very similar. When deciding which image looks the best from four almost identical images there is bound to be some randomness in the choices from the evaluation group. There were some clear images for the forest area as well but not as high percentage as for urban area. Therefore the choices for forest is a bit more evenly distributed than the remaining two but not as evenly as for the urban area. The results for these two areas, but especially urban area, are not as clear as for the other areas where there were a significantly smaller number of haze free images in the evaluation set. The fact that the distribution of choices by the group is rather even shows that none of the methods destroy an image if it was haze free. It is therefore a good outcome that the choices are evenly distributed which implies that there
were no significant change in the initially clear images.

The combination of the two methods performed really bad in desert, and this was confirmed by the evaluation where it was only chosen as the best result in 4% of the cases. This implies that the darker images, though mostly free from haze, had a color that made them unusable.

One reason for the original images in the island area to only being chosen in 3% of the cases, compared to 11%, 13% and 34% for the other areas, is that there were more images available in that set to be chosen from for the evaluation. This led to no initially clear images ending up in the evaluation set, hence the original image only being chosen as the best in 3% of the cases. The images used were randomly selected but the clear images were exchanged for another randomized image which had haze.

Using only the NIR method did not give any good results as previously discussed in section 4.3 and the evaluation group thought the same where the values in Table 4.1 ranks this method as the worst or second worst for all areas except for desert. Desert was also the only area in which any real difference could be seen.

The overall best performance came from using only dark channel. According to the evaluation group it gave the best result for desert and the second best result for the three other areas. Two of those areas, forest and islands, had the combination as the best setting and in both cases the NIR algorithm does not do any significant difference compared to only the dark channel. So to get a good result for as many areas as possible the best outcome will come from only using the dark channel algorithm.

The fact that the result becomes almost completely white after dark channel when there is nothing but sea in the image does not go with the rest of the results or the theory. When removing the haze the image should, and usually does, become darker since intensity is removed from the image. But in this case the intensity is increased from close to zero to the maximum value and therefore saturating the entire image. The cause for this is that there is not enough pixels that are not completely dark in the initial image for a good enough mask for the brightest pixels to be created, no pixels are marked as bright in said mask. Because of this the value for a can not be calculated, hence no transmission map and without them the result can not be calculated at all. The white image comes from the algorithm trying to produce a result and ending up with an image full of NAN which it tries to convert to something understandable, hence only white pixels. This only happens for images with no land and therefore a very small span in intensity that is placed in the absolute bottom of the scale. If these images could be detected beforehand these results would be easily avoided.
The result was expected to become slightly darker in most cases since the goal is to remove the haze and therefore some of the intensity in the image. However, the parts with no haze should not become much darker. This could come from a transmission map that is not perfect or from a wrong value for the atmospheric light. An incorrect transmission map leads to removing the wrong amount of haze, i.e. intensity, from the image. So a transmission map which has estimated it to be more haze than in reality would result in a darker image than there should be. If the atmospheric light is wrong the subtracted value is also wrong which makes the colors in the resulting image not conforming with reality. Therefore some of the darker results can originate in a combination of a transmission map that is slightly incorrect and a value for the atmospheric light that is untrue.

The clouds sometimes become blue or green when too much of the atmospheric light is removed from them. They are marked as really thick haze in the transmission map and therefore the algorithm tries to remove them. This results in subtracting an incorrect value in these regions and achieving a blue or green result. The regions that are not thick haze or cloud is handled in the correct way and not affected by this but the wrongly colored clouds destroy the total look of the images. If the results were to be used directly from the algorithm as they are they would not be very usable. But if the parts that held thicker haze and cloud were marked before the algorithm was applied those parts could be discarded after and only the clear and correctly dehazed parts can be selected and be used. Another option is marking those areas and not applying the algorithm there so the entire image will be usable.

By looking at the values presented in Table 4.2 - 4.5 one can see that they correspond to the observed look of the images and the results from the qualitative evaluation. From comparing the wanted values presented in section 4.1, which were large values for c and m and a small value on $\sigma$, with the results and the outcome from the qualitative evaluation one can see that they do not coincide in all cases and the values of the two "best" settings are very close to each other when they do not. This suggests that the quantitative evaluation can be seen as an initial guess for which method will perform best but that a qualitative evaluation has to be done as well.

These values clearly show that the combination of the methods gives the worst results for desert, which can be confirmed by looking at the images as well as from the qualitative analysis. The value for colorfulness is significantly lower for this setting than for the other ones which corresponds to this setting having a smaller color content than the other ones. This is exactly what can be seen when comparing the images depicting desert in Figure 4.1 - 4.5 where the result from the combination is much less vivid and is much darker than the other ones.

The value on $\sigma$ after dark channel and dark channel + NIR is larger for islands than for the other areas, which can be explained by the images depicting only water that becomes completely white. These images increase the total num-
ber of saturated pixels very much. If one were to omit these images from the evaluation these values would be closer to the corresponding values for the other areas. Since there is only a small number of white images compared to the total number of images in the island set, the mean value is not increased to a value that is completely off. One might not even look twice at this value if one did not know about the white images. These images might affect the other values as well but there is no clear difference in those values from the other areas.

5.2 Methods

When choosing Dark Channel Prior by He et al. one major factor in validating it was that it is one of the most referenced papers in the area of dehazing, and is used as a major reference in e.g. [16], [17] and [18].

Since the majority of the papers in the area of dehazing are intended for regular images it is not obvious how they will perform on remotely sensed images beforehand. The results in the papers looked really promising but the achieved results in this work did not turn out as well as hoped. One plausible explanation of this might be the use of regular images vs remotely sensed images.

5.2.1 Dark Channel

For the estimation of the atmospheric light a cloud mask is used in this work. This is to prevent the estimated color from becoming extremely far from the truth. The cloud masks used are not impeccable and therefore the color is estimated from some cloud pixels as well as clear pixels in most cases. This results in the cloud mask not being as important as first believed and stated in section 3.1 and the resulting images looks very similar regardless if the cloud mask is used or not. If the cloud mask was absolutely perfect for all images it might contribute to a better looking result. Estimating the atmospheric color from some cloud pixels leads to the color being lighter than what it would have been without the contribution from the cloud pixels. This can be a factor in that the result often is a bit darker than what is wanted.

This method removes some level of the atmospheric color in every pixel, even if there is no visible haze. This results in images darker than the original, even in the haze free regions. If this method is presented with a clear image it will calculate the atmospheric light and the transmission map and remove the "haze" which is not wanted. If there is no obvious haze the amount removed is very small so the resulting image is not that different from the original. The subtraction does however darken the image slightly, not by much, but it might affect the product the image is used in later. Since the transmission map determines how much haze is to be removed from each image individually not the same amount is removed from every image. This underpins the statement that the composition
of the image differs depending on the conditions in which the image is taken. Removing different amounts from each image results in the set to have a similar color across all images.

Since it is developed for regular images, in which the haze differs from the haze in satellite images, this method does not remove all the haze in satellite images. It does only remove the haze that is fairly thin, the thicker haze is too difficult for this method. This is mostly since there is not enough information about the scene behind those parts for the result to be any good.

5.2.2 NIR

This method did not work as well as expected. One probable explanation is that the Sobel filter that is used to detect the derivatives in the V-channel and the NIR-channel operates on pixel level so it detects really small changes in value. This results in derivative images in which every change is visible, down to the smallest ones between adjacent pixels. Since the source paper, [20], applies the algorithm to regular images that are a lot smaller than the satellite images used in this work a regular 3x3 Sobel filter works there. In this case where the images, and the objects in the images, are much larger, the 3x3 filter is too small and detects changes on the wrong scale. In the method using dark channel the difference in size between regular and satellite images was easily fixed by just increasing the kernel size of the guided filter from 15x15 to 50x50. Doing the same for the Sobel filter in this case would not work since increasing the size of the Sobel kernel is not equivalent with detecting changes over a larger area and ignoring the pixel wise changes. Using another filter to find the derivatives on a different scale could improve the results from this method. The output from these filters is used for determining the approximate images and if those images are more accurate and adapted to the scale of the satellite images, then the result would probably become better.

5.2.3 Evaluation methods

The lack of ground truth for this problem is a big inconvenience since the big part of the evaluation has to be subjective. This could have been a minor problem if the number of opinions on the results had been large but with the rather small group used here this evaluation is closer to subjective than objective. For a better, i.e. more objective, result from the qualitative evaluation a larger group probably should have been used. Now only ten people evaluated the images which is close to being too few. The results could still be analyzed but with more opinions the results might have been a bit clearer and easier to draw conclusions from.

The small number of available images for three of the areas forced some images that were clear from haze to begin with to being used in the qualitative evaluation. The different methods left those images almost unaltered which made it
hard for the evaluation group to decide which image looked the best and they were forced to pick one almost at random. This introduced an uncertainty in the study. For the optimal result the evaluation group should have been presented with many more images per area so the random choices would have been only a really small percentage of the total images. This would have meant that the study would have consisted of at least 400 different images (100 per area) instead of the 61 used now, which would have taken an unreasonable long time to go through and was not realistic to do. For the set depicting islands there were a larger number of images available but for the study a small number of images had to be chosen so the study would not be unreasonably long, as previously mentioned. The images chosen for islands were all hazy to begin with which explains the difference in the distribution of the choices between the island set and the other three.

From the explanation of the colorfulness metric in [22] a larger value implies a more colorful image. A hazy image is less colorful than a clear one and therefore this value should be as large as possible. The setting that performed best in the qualitative evaluation had the biggest increase in $m$ or were a close second in all cases. This implies that this metric is a good indicator of which setting gives the best result. However, it does not tell which setting is the best for certain and a qualitative analysis also has to be done.

The contrast metric can be very helpful in making sure that the contrast has not decreased in the process. However as long as it is positive, which corresponds to an increased contrast, its value does not say all that much about the quality of the image. If one were to simply look for the largest value for $c$ an unnatural-looking image with a high contrast might be chosen as the optimal result. Therefore this property is not that significant in the evaluation of the result as long as it is positive.

The percentage of saturated pixels is an important metric since a higher value means a result which has more pixels that does not look as they should. It could mean that a too big value has been set as the atmospheric color or that the transmission map is wrong. This value would be unnaturally large in the images that become completely white, therefore this metric is a good indicator that something has gone wrong in the process.
5.3 Wider Perspective

Most of the work in dehazing has been done on regular images and not on satellite images. The methods used and discussed in this work are good examples of this. Dehazing has come a long way regarding regular images but when searching for similar work on satellite images not much can be found. This work is meant as a contribution in that field so the satellite images can become better for e.g. a more accurate and easier visualization of the earth.
Conclusions

The method that works best overall is the one using the dark channel. This method removes most of the haze without making the result really dark as the combination can do. It removes thin haze almost completely in most scenes and if there is thicker haze it tries to remove it as well, with varying results. The blue residue from this thicker haze can be questioned if it really improves the image or if it only changes it to another bad state. For the usage of the images from this work it does not matter that much. This is since there are cloud masks which removes the parts of the image classified as cloud in the final use of the image. Therefore the final result does not rely on the color in those areas, whether they are white or blueish.

It is hard to say which method gives the result whose colors is closest to the reality since images from the surface of the earth on the corresponding locations was not available. The method that resulted in colors closest to the input image was the one using NIR, but this method barely removed any haze at all. Therefore the method that gave the best colors as well as removing a fair amount of haze was the one using the dark channel. The combination of the two methods gave good results in many cases but since it turned the desert really dark its result is not better than the one using only dark channel. Desert is a hard area for dehazing since there is not much contrast in this kind of images and they have a small span in intensity. Out of the four types of images evaluated in this work desert is by far the toughest one to get good results for. The other three areas give good results in most cases but not when the image has a small span in intensity as in the case with images covering only water.
The dark channel method works well if there is only thin haze in the image so the atmospheric light is estimated correctly or if there is a correct mask for not choosing white cloud pixels. If the atmospheric light is wrongly determined the colors in the resulting image is incorrect. This method can remove both homogeneous haze and streaks of haze as long as there is enough information about the land below in the image. In most cases clouds are left unaltered but, as mentioned, if the algorithm tries to remove too much intensity in those areas the clouds can become blueish.

### 6.1 Future Work

Some of the resulting images become a lot darker than their input image, so by comparing e.g. the mean intensity of the parts of the image not classified as cloud and applying a brightening algorithm on the images whose mean has decreased more than wanted the results could become better. Right now some seemingly good results are too dark so they can not be used in a proper way but by brightening them they could be used and possibly even be better than the images chosen as the best now.

Investigating how the dark channel looks in different conditions, e.g. no haze, moderate haze or haze with additional clouds, could improve the performance on haze free images. This algorithm is supposed to run on every image without a human deciding if there is haze or not beforehand. Therefore it is important that the color of images with only a small amount of haze does not change too severely when run through the algorithm. If there is no haze or clouds in the image its dark channel will consist of only relatively low values. So by e.g. setting a threshold for the mean value of the dark channel or the number of pixels with a value above a threshold, a decision can be made to run the algorithm on the current image or not. I.e. if the mean is below a threshold it would mean that there is not enough haze or clouds in the image for the algorithm to be run on this particular image. This would however not take the opaque clouds into consideration since they would increase the mean a lot, so images with only clouds but no haze would still pass as having to be run through the algorithm, if the clouds are not excluded by using a cloud mask.

The dark channel method works well on most of the scenes covered in this work. But for it to properly work on all images and scenes, more difficult scenarios has to be investigated e.g. snowy scenes. Thus investigating scenarios that are not mentioned in this work could improve the overall performance for haze removal from all images, regardless of what they depict.
Appendix
Different versions of the same image are shown side by side here, for an easier comparison between the results from the different methods. All images have been shown in section 4 but there they are grouped by method rather than image.
Figure A.1: Desert
Figure A.2: Forest
Figure A.3: Urban Area
Figure A.4: Island
Figure A.5: Island
Figure A.6: Island


