

MONITORING AFFORESTATION WITH VERY HIGH RESOLUTION SATELLITE IMAGERY IN KENYA

METHODS FOR INDIVIDUAL TREE COUNTING

Ellen Ghyselbrecht

Studentennummer: 0160761

Promotor(en): prof. dr. ir. Fieke Vancoillie

Masterproef voorgelegd voor het behalen van de graad master in de Bioscience engineering.

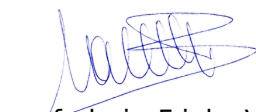
Academiejaar: 2022 - 2023

De auteur en promotor geven de toelating deze scriptie voor consultatie beschikbaar te stellen en delen ervan te kopiëren voor persoonlijk gebruik. Elk ander gebruik valt onder de beperkingen van het auteursrecht, in het bijzonder met betrekking tot de verplichting uitdrukkelijk de bron te vermelden bij het aanhalen van resultaten uit deze scriptie.

The author and promoter give the permission to use this thesis for consultation and to copy parts of it for personal use. Every other use is subject to the copyright laws, more specifically the source must be extensively specified when using results from this thesis.

Gent, January 17, 2023

The promotor,

A handwritten signature in blue ink, appearing to read 'Fieke Vancoillie', with a stylized flourish at the end.

prof. dr. ir. Fieke Vancoillie

The author,

A handwritten signature in blue ink, appearing to read 'Ellen Ghyselbrecht', with a large, sweeping flourish at the end.

Ellen Ghyselbrecht

ACKNOWLEDGEMENT

First of all, I want to thank my promotor prof. dr. ir. Fieke Vancoillie and my tutor ir. Jasper Feyen for the coaching and for granting me extra time to finish my master's dissertation. I also want to thank Jan and Samuel of Better Globe Forestry for the opportunity of working on this project and for allowing me to visit the Kiambere plantation. Furthermore, I want to thank Halima and Kelvin for carrying out the largest part of the data collection.

Besides that, I want to thank Eva and Tanguy for supporting me through this thesis and for the encouraging words during this difficult past year. I also want to say thanks to my roommates: my sister Evi, Watson the dog and Bloempje the cat for the necessary distraction. Lastly, I want to thank my parents for supporting me through university.

CONTENTS

Acknowledgement	i
Content	ii
Samenvatting	iii
Abstract	iv
1 Introduction	1
2 Literature review	3
2.1 Satellite imagery	5
2.2 Image preprocessing and enhancement	6
2.2.1 Pan-sharpening	6
2.2.2 Band selection	7
2.2.3 Filters	7
2.3 Template matching	9
2.3.1 Templates	10
2.3.2 Similarity measure	14
2.4 Local maximum filtering	15
3 Material and methods	18
3.1 Area description	18
3.2 Data collection	21
3.2.1 Satellite imagery	21
3.2.2 Field survey	22
3.3 Regions of interest	24
3.3.1 Cloud mask	24
3.3.2 Tree mask	24
3.4 Template matching	25
3.4.1 Parameter optimisation	26
3.5 Local maximum filtering	31
3.6 Accuracy assessment	32
4 Results and discussion	34
4.1 Template matching	34
4.1.1 Tree parameters	34
4.1.2 Farms	36
4.1.3 Plantation	40
4.2 Local maximum filtering	45
4.3 Method comparison	47
4.4 Recommendations	48
4.4.1 Data collection	48
4.4.2 Algorithm optimisation	48
4.4.3 Satellite and aerial imagery	49
5 Conclusion	51
Bibliography	53
Appendix A Locations of the data collection	59
Appendix B Descriptive statistics	61
Appendix C Template library	62
C.1 Farms	63
C.2 Plantation	72

SAMENVATTING

Better Globe Forestry (BGF) is een bedrijf gevestigd in Kenia dat in 2004 is opgericht met als missie om door middel van sociaal ondernemerschap zoveel bomen te planten als er mensen op deze planeet zijn. Nu wil BGF de bebossingsinspanningen monitoren die door de jaren heen zijn geleverd en zien hoeveel bomen er nog leven, en vooral, hoeveel *Melia volkensii* bomen. Omwille van het enorme gebied dat moet worden gemonitord, is dit alleen mogelijk met behulp van teledetectie. Daarom is een verkennend onderzoek de eerste stap om tot een duurzaam monitoringinstrument te komen. In deze masterproef werd onderzocht of het mogelijk is *Melia volkensii* bomen nauwkeurig te tellen in beide bebossing layouts.

Tijdens een veldcampagne tussen oktober 2021 en januari 2022 werden boomgegevens verzameld in de regio van Kiambere in beide bebossing layouts. Verder werd gebruik gemaakt van beelden met een zeer hoge resolutie verkregen door de pléiades-satelliet, met een panchromatische band met een spatiale resolutie van 0.5 m en drie multispectrale (MS) banden (d.w.z. rood, groen, blauw en nabij-infrarood) met een ruimtelijke resolutie van 2 m. Met behulp van het satellietbeeld en de verzamelde gegevens werden twee methodes getest: een op template matching (TM) gebaseerde methode en een op lokale maximum filter (LMF) gebaseerde methode. Met behulp van empirische en synthetische templates werd de TM methode eerst geoptimaliseerd met behulp van de vier MS-banden en de brightness feature. In de boerderijen en de plantage werden de beste performance (F-score = 0.411 en plot level accuracy = 0.787) verkregen met behulp van de brightness feature met drie empirische templates. Voordat de tweede methode werd toegepast, werden mathematical morphology procedures gebruikt om het beeld te filteren. Deze methode presteert ondermaats in vergelijking met de TM-methode met een F-score van 0.023 met behulp van de groene band in de boerderijen en een plot level accuracy van 3.425 met behulp van de brightness feature in de plantage. De positionele nauwkeurigheid (3 m) van de bomen was het meest uitdagende probleem tijdens de trainingsfase van het TM - algoritme en de validatiefase van beide methoden. Om de bebossingsinspanningen van BGF nauwkeurig te monitoren, zijn meer ruimtelijk nauwkeurige boomposities nodig, moet het TM-algoritme verder worden geoptimaliseerd en zijn beelden met een hogere spatiale resolutie nodig.

ABSTRACT

Better Globe Forestry (BGF) is a Kenyan-based company founded in 2004 with a mission to plant as many trees as there are people on this planet through social entrepreneurship (BGF, nd). Now BGF wants to monitor the afforestation efforts that have been made throughout the years and to see how many trees are still standing and, in particular, how many *Melia volkensii* trees. Due to the vast area that needs to be monitored, this is only possible with remote sensing. Therefore, an explorative study is the first step in obtaining a sustainable monitoring tool. In this master's dissertation, it was researched if it is possible to accurately count *Melia volkensii* trees in both layouts.

During a field survey between October 2021 and January 2022, tree data was collected in the region of Kiambere in both afforestation layouts. Furthermore, very high-resolution (VHR) imagery acquired by the pléiades satellite was used, containing a panchromatic band with a spatial resolution of 0.5 m and three multispectral (MS) bands (i.e., red, green, blue and near-infrared) with a spatial resolution of 2 m. Using the satellite image and the collected data, two methods were tested: a template matching (TM) based method and a local maximum filter (LMF) based method. The TM method was first optimised using the four MS bands and a brightness feature using empirical and synthetic templates. In the farms and the plantation, the best performance (F-score = 0.411 and plot level accuracy = 0.787) was obtained using the brightness feature with three empirical templates. Before applying the second method, mathematical morphology procedures were used to filter the image. This method underperforms compared to the TM method with an F-score of 0.023 using the green band in the farms and a plot level accuracy of 3.425 using the brightness feature in the plantation. The positional accuracy (3 m) of the trees was the most challenging problem during the training stage of the TM - algorithm and the validation stage of both methods. To accurately monitor the afforestation efforts of Better Globe Forestry, more spatial accurate tree positions are needed, the TM algorithm needs further optimisation and higher resolution imagery is needed.

1. INTRODUCTION

Better Globe Forestry (BGF) is a Kenyan-based company founded in 2004 with a mission to plant as many trees as there are people on this planet through social entrepreneurship (BGF, nd). By planting these trees, the company wants to finance and sustainably implement its vision to eradicate poverty and corruption in Africa. In 2009, BGF expanded their operations to Uganda. The company is firmly embedded in East African society and tries to work closely with local communities.

BGF's expertise lies in the commercial afforestation of drylands in East Africa, particularly with the indigenous mahogany species, *Melia volkensii*, in an agroforestry layout. Other species, such as *Melia azedarach*, are also being planted. Trees in the agroforestry layout are planted with wide spacing to allow farmers to intercrop with food crops. Apart from this first layout, there are classic plantations, the Kiambere model plantation and the Nyongoro plantation. Now BGF wants to monitor the afforestation efforts that have been made throughout the years and to see how many trees are still standing and, in particular, how many *Melia volkensii* trees. Due to the vast area that needs to be monitored, this is only possible with remote sensing. Therefore, an explorative study is the first step in obtaining a sustainable monitoring tool. The development of a prototype monitoring tool contains two objectives:

1. Counting of individual tree crowns in both afforestation layouts.
2. Differentiating *Melia volkensii* from other tree species.

The first objective is the subject of this master's dissertation. The second objective is the subject of the master's dissertation of Forceville (2022), which will be often referred to throughout this thesis.

Many individual tree counting methods have been developed (Larsen et al., 2011; Zhen et al., 2016; Cheng and Han, 2016; Ke and Quackenbush, 2011; Erikson and Olofsson, 2005). However, this master's dissertation will solely focus on two approaches: template matching (TM) and local maximum filter (LMF). The methods are developed based on a study area in Kiambere where *Melias* are planted in two afforestation layouts: a plantation and an agroforestry layout. In the following text, the model plantation in Kiambere will be referred to as 'the plantation' and the agroforestry systems as 'farms'.

The research questions this master's dissertation will try to answer are the following:

1. *Is it possible to accurately count Melia Volkensii trees using very high resolution (VHR) satellite imagery?*
2. *Which method has the highest performance?*

3. Is there a different method needed for the different afforestation layouts?

Considering the limitations of the Pléiades satellite image with a spatial resolution of 0.5 m, it is expected that the trees with a crown area smaller than 0.25 m² will go undetected.

In chapter 2, literature about individual tree crown detection (ITCD) is reviewed. The methods and materials are elaborated in chapter 3 followed by a results and discussion section (chapter 4). Lastly, in chapter 5, an overall conclusion about the study is made.

2. LITERATURE REVIEW

Researchers have developed several automatic or semi-automatic methods for individual tree crown detection with the use of very high-resolution (VHR) images (Ke and Quackenbush, 2011). For different satellite imagery, the interaction between the forest canopy and the incident light can result in various spatial and spectral characteristics depending on the site, sensor type, image scale and timing of the image collection (Lamar et al., 2005). These characteristics (or a combination of characteristics) help to assist the separation of individual tree crowns, e.g., local minima reflectance values due to shadow between individual crowns (Culvenor, 2002; Gougeon, 1995), local maxima reflectance values representing the tree crown top (Culvenor, 2002; Pouliot et al., 2002) and crown shape characteristics (Pollock, 1996). The approaches published in the literature are usually developed and evaluated for one specific sensor and one specific forest (Larsen et al., 2011). This makes it hard to draw an objective conclusion about what method performs best for a given situation. Several factors decide the performance of a method; for example, a man-made forest is structurally very simple, while an unevenly-aged naturally regenerated forest can be structurally complex, making ITCD hard. Larsen et al. (2011) mention that if the relative performance of different tree detection methods on different forest types was available, an expert system could, for a specific image, choose the best-suited method. In a review study conducted by Zhen et al. (2016), it was demonstrated that most of the methods developed between 1990 – 2015 for tree detection were developed for closed softwood forests (i.e. deciduous forests) and a smaller amount for closed hardwood forests (deciduous stand). Moreover, most studies yield higher accuracy in conifer stands than in deciduous stands and in sparse rather than dense canopies (Zhen et al., 2016).

Different authors divide ITCD methods into different groups. Larsen et al. (2011) have divided the methods available for passive monoscopic imagery into six groups, namely: local maximum filtering (LM) (Wulder et al., 2000; Pouliot et al., 2002; Wang et al., 2004), valley following (VF), template matching (TM) (Pollock, 1996), scale-space (SS) theory-based methods (Brandtberg and Walter, 1998; Lindeberg, 1996), Markov random fields (MRFs) (Maillard and Gomes, 2016) - based methods and marked point process (MPP) methods (Gomes et al., 2018). The simplest approach for tree crown detection is based on finding local maxima on the original or a pre-processed grey-scale image, which are assumed to be tree crowns. VF methods also make the assumption that tree tops are the brightest spots while all the areas between the crowns appear darker due to shading. These shaded pixels are chained together, forming a boundary of tree crowns (Gougeon, 1995). The third group, template matching, uses templates or models to detect trees in the image (Pollock, 1996). SS methods refer to methods that try to locate trees of different sizes in the same grey-scale image using different scale levels (Larsen et al., 2011). Ke and Quackenbush (2011) has stated that scale was a critical factor influencing tree detection accuracy. For example: for images where tree crown sizes are variant and the image resolution is invari-

ant, it is difficult to detect all crowns simultaneously: tree crowns that are small compared to the image resolution may not be detected and large crowns may be identified as multiple trees. MRF and MPP are two probabilistic methods (Descombes and Pechersky, 2006).

Maillard and Gomes (2016); Ke and Quackenbush (2011) have discriminated two more additional segmentation methods, namely: watershed segmentation (WSS) and region growing (RG). Watershed segmentation is a method that views a greyscale image as a topographic surface where each pixel's grey level intensity equals the elevation (Biswas et al., 2020). Firstly, the local maxima, interpreted as tree crowns, are inverted to local minima, forming valleys. Next, the image is slowly flooded and water accumulates in the valleys until it overflows into adjacent valleys. Lastly, dams are built on the watershed lines to prevent overflowing from one basin into another. These dams form the segments and delineate tree crowns. RG is another segmentation technique that groups pixels based on predefined criteria such as specific intervals of intensity or colour (Ke and Quackenbush, 2011). Starting from seed pixels (random or based on prior knowledge), the neighbouring pixels are examined and added to the growth region if their properties meet the criteria (Gonzales 2008).

It should be noted that many approaches use hybrid methods for the detection and delineation of tree crowns Maillard and Gomes (2016); for example, the markers (i.e. tree tops) obtained with local maxima filter are often used as a subsequent step for tree crown delineation using VF, WSS or RG (Pouliot et al., 2002; Erikson, 2004, 2003, 2006; Wang et al., 2004). Gomes and Maillard (2013) have also proposed a hybrid method integrating TM and MPP for detecting urban trees and trees in orchards in Brazil, France and Italy using VHR imagery.

Some methods extend to multiview images as well, e.g. Gong et al. (2002); Kempf et al. (2021); Korpela et al. (2007) have used a model-based approach in 3D using stereo-pairs. Hirschmugl et al. (2007); Heipke et al. (2007); Xiao et al. (2018) have obtained a canopy height model (CHM) from stereoscopic images and performed LMF or WSS on this 3D CHM. However, in recent years, high-density LiDAR has become the remote sensing data source of choice to directly obtain a CHM for tree crown detection and delineation (Larsen et al., 2011). The obtained CHM is then often processed using algorithms that are closely related to the methods used for passive optical imagery (Larsen et al., 2011), such as LMF and methods based on SS theory (Brandtberg et al., 2003). Several studies combine LiDAR (active optical data) and MS passive optical data for detection and segmentation (Heipke et al., 2007). Larsen et al. (2011); Zhen et al. (2016) have mentioned that even though 3D data has become more available through LiDAR, the development of methods for passive optical images is still justified because VHR satellites may be used to provide data where airborne laser scanning is too expensive. However, Zhen et al. (2016) has mentioned that by integrating MS data and LiDAR data, commission and omission errors may be reduced.

In the approaches mentioned above, some detect tree location before delineating crowns, while others combine these fundamental components in a single step (Ke and Quackenbush, 2011). According to Ke and Quackenbush (2011), some even consider tree detection as equivalent to tree delineation. Only TM, MRF, MPP, LMF and SS are tree detection methods (Kempf et al., 2021; Ke and Quackenbush, 2011).

Some algorithms work well in relatively closed canopies, while others are more suited to individual trees with little crown overlap (Pouliot et al., 2002). Larsen et al. (2011) have stated that because TM can use shadows cast by trees, the method is more suitable for open forests, while VF and LM-based methods are more suited for denser areas where trees are separated by areas of shade. Larsen et al. (2011) also confirmed this statement by comparing the methods in different images with different crown closures (from a homogeneous plantation, an area with isolated tree crowns, to an extremely dense deciduous forest). Gougeon (1997) has used a locally adaptive method that applies the LMF in areas identified as dense and a modified method in areas identified as open using. The dense and sparse areas are identified based on a gradient image.

Since the objective of this study is to count trees in a dense (plantation) and sparse (agro-forestry) layout, template matching and local maximum filtering will be further elaborated and tested in this study. LMF is the simplest method and has shown promising results in dense forest stands, while template matching is more complex but has shown promising results in open stands. Before applying the algorithms, the correct remote sensing data must be selected (section 2.1). Moreover, in further research, these methods could also be integrated into more complex tree crown detection and delineation algorithms, as mentioned above. Once the data has been obtained, several preprocessing and enhancement steps can be applied to the data (section 2.2). The different tree crown detection methods will be discussed in section 2.3 and 2.4.

2.1 Satellite imagery

Tree crowns in VHR show within-crown brightness variation due to the effect of branches and branch shadow patterns. According to Pouliot et al. (2002), this affects all tree detection algorithms. This effect can be reduced by decreasing the image resolution and thus averaging the spectral data over a larger areal unit. On the contrary, crown boundaries become less distinct at lower image resolutions, making them harder to identify. For this reason, Pouliot et al. (2002) suggests using a one-dimensional ratio of the average crown diameter to the pixel size to characterise the crown shape at different image resolutions. A ratio between 3:1 and 19:1 is proposed as lower ratios do not retain distinct crown boundaries and higher ratios may contain too much within-crown variation. The ratio optimum will differ depending on the algorithm's sensitivity to this within-crown brightness variation and boundary brightness gradient.

Table 2.1 gives an overview of the available satellites and their respective spatial resolution and spectral bands. The increase in spatial resolution was not always accompanied by an increase in spectral resolution for VHR, which is often restricted to the panchromatic band: there is only one available satellite with VHR in multispectral bands (at this time) (Maillard and Gomes, 2016).

Kaszta et al. (2016) has mentioned that, in African savanna ecosystems, it is also important to consider seasonality while classifying land cover components, due to the huge contrast in vegetation state during the wet and dry periods (Venter et al., 2003). During the rainy season, the image is affected by dense cloud cover and a high amount of atmospheric water vapour leading to large gaps in the image (Kaszta et al., 2016). Nonetheless, images

2. Literature review

Table 2.1: Available very high-resolution satellites with their launch year, spatial resolution (m) and spectral bands i.e.: red (R), green (G), blue (B), near-infrared (NIR), near-infrared 2 (NIR2), coastal (C), yellow (Y) and red edge (RE)(Maillard and Gomes, 2016).

Satellite	Launch year	PAN resolution (m)	MS resolution (m)	MS bands
Ikonos II	1999	0.82	3.2	B, G, R, NIR
Quickbird-2	2001	0.65	2.62	B, Gr, R, NIR
WorldView-1	2007	0.46	/	/
Geoeye-1	2008	0.46	1.84	B, G, R, NIR
WorldView-2	2009	0.46	1.85	C, B, G, Y, R, RE, NIR, NIR2
Pléiades 1A	2011	0.50	2.00	B, G, R, NIR
Pléiades 1B	2012	0.50	2.00	B, G, R, NIR
Kompsat-3	2012	0.70	2.80	B, G, R, NIR
SkySat-1	2013	0.90	2.00	B, G, R, NIR
WorldView-3	2014	0.31	1.24	C, B, G, Y, R, RE, NIR, NIR2
SkySat-2	2014	0.90	2.00	B, G, R, NIR
Kompsat-3A	2015	0.55	2.20	B, G, R, NIR
WorldView-4	2016	0.30	1.24	R, G, B, NIR

during the wet season the images are often preferred because they show well-developed vegetation cover and leaf-on conditions (Liu et al., 2015). The high photosynthesis rate can also lead to confusion between different vegetation types (i.e. trees, shrubs, grass and crops) (Kaszta et al., 2016). On the contrary, dry-season images are cloud-free but it might be difficult to detect leafless deciduous trees.

2.2 Image preprocessing and enhancement

The goal of preprocessing in the context of ITCD is similar to most remote-sensing applications. Most commonly this includes the removal of noise, enhancing the distinction between the objects (i.e. tree crowns) and the background and masking irrelevant areas (Ke and Quackenbush, 2011). For VHR imagery this often includes the selection or the derivation of the most appropriate image band(s). Different preprocessing techniques will be discussed in section 2.2.1-2.2.3, in no particular order. The order in which enhancement and preprocessing techniques are applied depends on the used method.

2.2.1 Pan-sharpening

Most of the operating earth observation satellites such as Landsat, Spot, Ikonos and Pléiades provide panchromatic images at a higher spatial resolution than their multispectral

mode (Ehlers et al., 2010), which provides high spectral resolution (Ozendi et al., 2016). The resolution ratio (RR) measures the difference in spatial resolution between the panchromatic and multispectral modes and is calculated as the ratio of the ground sampling distances (Klonus and Ehlers, 2009). In order to generate MS images with the same high spatial resolution as the PAN images, pan-sharpening methods are developed (Ozendi et al., 2016) by combining both images (Ehlers et al., 2010). Ehlers et al. (2010) have stated that most of these methods work well with images that were acquired at the same time by one sensor (single-sensor, single-date fusion). Some of the methods are the Wavelet Transform (WT), the Brovey transformation, intensity-hue-saturation (IHS) transformation, principal component analysis (PCA) sharpening, the high pass filtering (HPF), fusion and the Gram Schmidt (GS) fusion. Each pan-sharpening method has both advantages and disadvantages: while some focus on enhanced spatial information, others focus more on improved colour information. Alimuddin et al. (2012) have concluded that GS and PCA have shown better results for colour reproduction compared to HPF and W-IHS methods when applied to WorldView-2 imagery. Pu et al. (2018) have used the GS spectral sharpening approach for the classification of trees with the use of Pléiades imagery.

2.2.2 Band selection

Salamí et al. (2019) have expressed that RGB images are not the most convenient for the soil/plant classification problem. The NIR band is a popular choice to detect tree crowns, because of its sensitivity to vegetation differences (Ke and Quackenbush, 2011). Quackenbush et al. (2000) compared the results when using the average different red bands and the square root of the sum of these bands but found little difference in the results in comparison to using the NIR band alone. However, other studies exist where other bands or combinations are used. Wang et al. (2004) used a PCA and only used the first component image, containing the most variation, for the classification of spruce trees in British Columbia (Canada). Bunting and Lucas (2006) have suggested that the ratio between a red edge band and a red band results in better treetop detection and boundary edge definition. Pouliot et al. (2002) have proposed a method where the absolute difference between NIR and red spectral bands is used to reduce the spectral response from soil and non-apex crown pixels in VHR images. Other studies (Brandtberg and Walter, 1998; Khan and Gupta, 2018) transformed RGB colour images into intensity-hue-saturation (IHS) to suppress non-forest areas. Salamí et al. (2019) have transformed an RGB image captured with a drone (DJI Phantom4) into CIE Lab colour space to detect olive trees in Spain.

2.2.3 Filters

Smoothing filters are used for blurring and noise reduction to get a gross representation of the objects of interest, such that the intensity of smaller objects blurs in with the background (Gonzalez and Woods, 2008). Different filters can be used for image smoothing. The mean and the Gaussian filter are some of the most known and straightforward. The Gaussian filter is often preferred as it preserves edge features better than a mean filter (Ke and Quackenbush, 2011). Image smoothing can also negatively affect working with a coarser image resolution by blurring the edge between the crown area and the background.

2. Literature review

Mathematical (MM) is a tool widely used in diverse image-processing tasks. In remote sensing, it is used to generate morphological features for the discrimination of objects in high spatial images. Without going into the mathematical background of MM, the basic operations and their effect on images will be explained. The two fundamental operators are erosion and dilation (Kalapala, 2014). These operators are applied using a structuring element (SE), which can have different shapes (e.g. disk, square, lines, ...) and sizes, depending on the case study. The first row of figure 2.1 shows examples of SE with the centre indicated by a dot (Gonzalez and Woods, 2008). The second row of figure 2.1 shows how these SE are converted into rectangles since this is required when working with images (Gonzalez and Woods, 2008). Erosion shrinks objects in the image, while dilation thickens objects in the image.

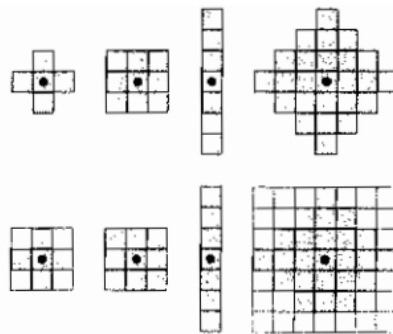


Figure 2.1: The first row shows examples of structuring elements (SE) and the second row shows the converted SE into rectangles by adding background pixels (Gonzalez and Woods, 2008). The black dots represent the centre of the SE.

Figure 2.2b demonstrates the erosion of a binary image (one/grey pixels represent the object, zero/white pixels the background) with square 3x3 SE with the origin in the centre. For each pixel in the image, if the SE is completely contained by the object in the image, the pixel is left unchanged, or else it is deleted (Gonzalez and Woods, 2008; Soille et al., 1999). Figure 2.2c demonstrates the dilation of the same image with a square 3x3 SE with the origin in the centre. For each pixel in the image that has a value of one at the origin, all pixels within the SE are changed to 1 and else the pixels are left unchanged. These basic operations can also be applied to greyscale images. When performing an erosion on a greyscale image, the value is replaced with the minimum value within the SE instead of deleted, resulting in a darker image with bright features reduced (Gonzalez and Woods, 2008). Conversely, when performing a dilation on a greyscale image, the values are replaced with the maximum value within the SE, resulting in a brighter image and reduced dark features.

By combining these basic operators (i.e. erosion and dilatation), other operators, such as opening and closing, can be derived. An opening is an erosion (shrinking) followed by dilation (regaining shape), while closing is a dilation followed by erosion (Khan and Gupta, 2018). These morphological filters are non-linear transformations that locally modify the geometry of features in the image (Terol-Villalobos and Vargas-Vázquez, 2005). Opening of the image results in a new image where the small, bright object (compared to the SE) is deleted and replaced with a greyscale value of their surroundings Khan and Gupta (2018). Closing of an image does the opposite and removes small dark pixels, while bright pixels

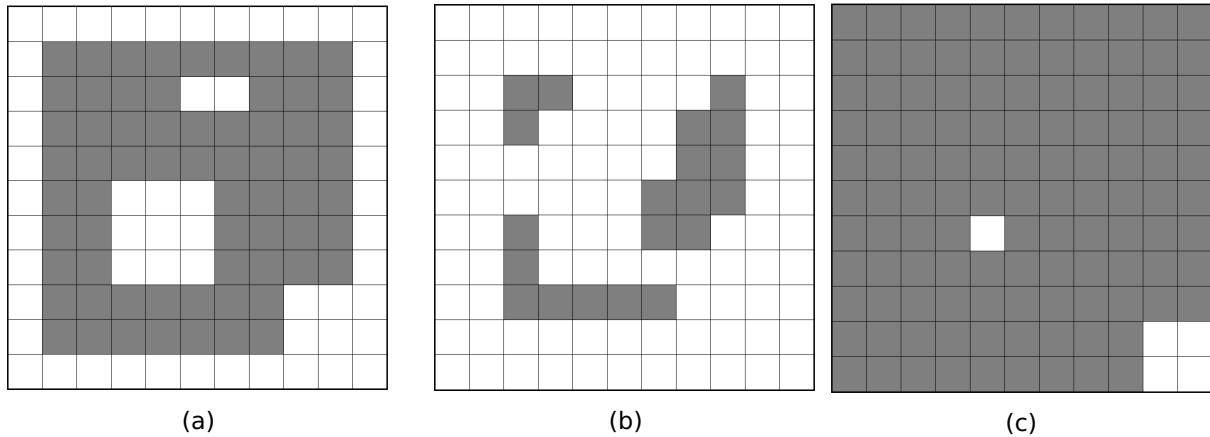


Figure 2.2: Erosion (b) and dilation (c) of the image (a) with a 3x3 square structuring element. White pixels represent the background/ zero pixels and grey pixels represent the object/ one pixel.

are left unchanged. Benediktsson et al. (2003) has mentioned that it is a common practice to use the opening and closing transform in order to isolate bright and dark structures in images. Opening and closing both generally smoothen the contour of an object Gonzalez and Woods (2008). Depending on the interaction between the SE and size and the size of the structure, some structures may have a high response for a given SE (Benediktsson et al., 2003). When the exact size of the structures (i.e. trees) is known, the appropriate SE can be chosen. Biswas et al. (2020) have tested the impact of the iterative opening of the image with SE with different kernel sizes to detect mangrove patches and have found that increasing the number of iterations decreased the overall accuracy of tree detection. The study concluded that one iteration of morphological opening promoted the overall accuracy when a square SE of 3x3 was used since additional openings remove not only noisy pixels but also valid tree pixels (Biswas et al., 2020). However, they have also mentioned that these parameters may change depending on lighting conditions and the heterogeneity of the vegetation matrix in which the trees are embedded. Salamí et al. (2019) have used consecutive opening and closing operations prior to applying a template matching algorithm (section 2.3) to detect olive trees in Spain.

Opening by reconstruction is an erosion followed by morphological reconstruction while closing by reconstruction is a dilation followed by morphological reconstruction (Khan and Gupta, 2018). Both these reconstruction operations partially restore the objects after erosion or dilation has been applied. Kalapala (2014) has mentioned that reconstruction-based opening and closing are more effective than the standard opening and closing at removing noise without affecting the overall shape of the objects. Kalapala (2014) has used opening and closing by reconstruction for tree counting on a plantation and in an urban area.

2.3 Template matching

Template matching is an object detection method which can be used to detect trees in remotely sensed imagery and was originally presented by Pollock (1996). Hung et al. (2012) has stated that the performance of object detection algorithms purely based on statistical

information is limited by the quality of the training data. Therefore, the use of an algorithm based on a tree model (template) is proposed, which incorporates both the object shape and context information. As shown in figure 2.3 this approach has two main steps. First, a template (T) needs to be generated with the use of a model or by learning from a training data set (Cheng and Han, 2016). The second step is calculating the similarity between a given source image and the templates at each possible location to find the best matches according to a minimum distortion or maximum correlation measure (threshold). Lastly, a post-processing step can be applied to remove hits that are too close to each other and probably originate from the same tree by implementing a minimum distance or maximum overlap parameter (Erikson and Olofsson, 2005). The first two steps will be explained in more detail in sections 2.3.1 and 2.3.2.

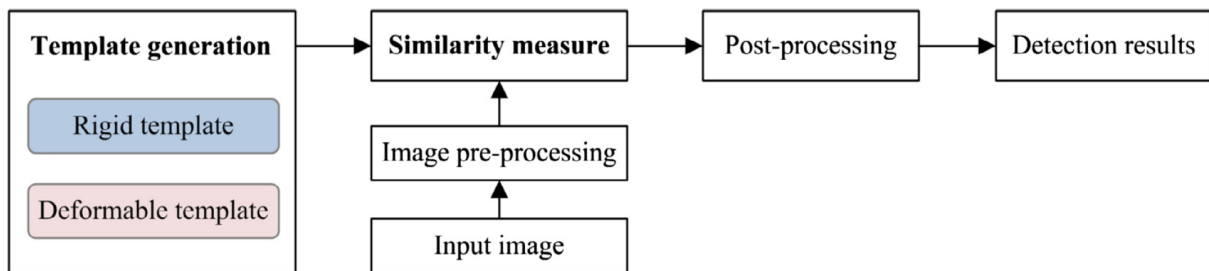


Figure 2.3: Template matching flowchart as suggested by Cheng and Han (2016).

2.3.1 Templates

TM-based methods include generating a series of models (or templates) that characterize the trees at different locations on the image. Thus taking both the trees' geometric and radiometric properties into consideration (Wang et al., 2004). A first and more traditional approach to template matching is to use a comprehensive model based on representative examples of the object of interest (Pratt, 1991). This example-based approach is attractive because of its much simpler implementation than the synthetic approach: there is no need for a detailed understanding of the process through which the trees in the three-dimensional world become two-dimensional images (Pollock, 1996). Pollock (1996) has mentioned that if there is a large natural spatial (e.g., shape, size and orientation of the object) and pattern (e.g., foliage density) variation within object images, the manual effort to collect a representative set of examples may be impractical. Alternatively, a synthetic comprehensive image model can be produced. This strategy permits readily acquired knowledge of the sensing system, the sensing situation and physical properties of individual tree crowns in the scene to be used, making it less reliant on specific scene and sensing conditions than other tree crown recognition procedures (Pollock, 1996).

Several different types of synthetic templates for object detection can be distinguished. Cheng and Han (2016) make the distinction between two groups: rigid and deformable TM. Rigid templates can be effective in some applications but have several disadvantages. The template must be precise, making it sensitive to shape and density variation. An exact geometric template is often unavailable because of changes in viewpoint and large intra-class variations among objects (e.g. trees). Deformable templates are more powerful in dealing with shape deformations and intra-class variations. It imposes geometrical constraints on

the shape and integrates local image evidence. Deformable templates can once more be divided into two classes: free-form and parametric deformable templates (Jain et al., 1998). Free-form templates represent an arbitrary object shape by constraining, e.g. continuity and smoothness. The most popular models are the active contour models, also known as the snake models. The snake is modelled as able to deform elastically, but deformation increases its internal energy causing a restitution force that tries to bring it back to its original shape (Jain et al., 1998). In parametric deformable templates, its variation is constrained by a parametric formula or by using a prototype and deformation modes. The latter is often used when some prior information about the geometrical shape (e.g. tree crown shape) is available.

Templates generated for tree detection can be considered parametric deformable templates constrained by a parametric formula. The basic elements to create a synthetic tree crown image model are prior knowledge of the tree shape, the position of the light source (sun) and the position of the camera (satellite) (Larsen, 1997; Pollock, 1996). Based on these elements, a template can be generated, as seen by the satellite. As mentioned at the beginning of this section, the trees' geometric and radiometric properties are taken into consideration. Both these aspects will be further elaborated.

Geometric aspects

It is not possible to generalise and model the exact shape of a natural object (Hung et al., 2012). Therefore a simple geometric model is used to approximate the outline of the tree crown. The three-dimensional shape outline can be described by a shape function, based on prior knowledge of the target object. Pollock (1996) has proposed the use of a generalised ellipsoid of revolution (GER) to represent a crown given by equation 2.1.

$$\frac{(z^2)^{n/2}}{r_3^n} + \frac{(x^2 + y^2)^{n/2}}{r_1^n} = 1 \quad (2.1)$$

The tree crown coordinates (x,y,z) are defined in a Cartesian coordinate frame, such that the unit vector z points vertically upwards and the unit vector x points toward the true north. The parameter n is a positive, non-zero shape factor that determines the curve of the ellipsoid (figure 2.4). When n decreases past 1 to 0, the curve becomes increasingly upward concave, while if n increases from 1, the curve becomes increasingly downward concave. The equation assumes that the tree crowns are rotationally symmetric about a vertical axis (Erikson, 2004). In other words, the radius of the width of the crown on the x -axis (r_1) is equal to the one on the y -axis (r_2). The modelled crown volume is bound by the xy -plane (i.e. z is only positive). This model has been used by Larsen (1997) and Larsen and Rudemo (1998) to detect Norway spruce (*Picea abies*) and by Erikson and Olofsson (2005) to detect pines (*Pinus sylvestris* L.). Gomes et al. (2018) has represented tree crowns by a half-dome with a fixed diameter to crown height ratio, but varying radii. This is to detect a wide diversity of tree species in Belo Horizonte (Brazil), Sydney (Australia) and California (USA) in a scattered layout, using a template matching - marked point process (TM-MPP) approach.

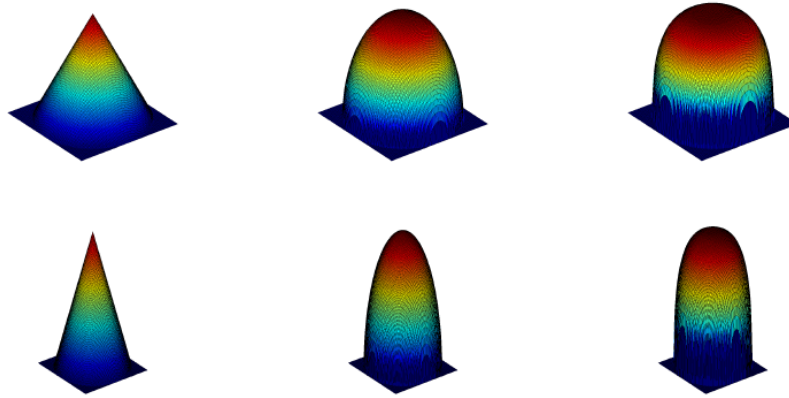


Figure 2.4: Examples of the generalised ellipsoid generated by formula 2.1 for $r_1 = 50$ and varying r_3 and n . In the upper row $r_3 = 100$ and in the lower row $r_3 = 200$. In the left column $n=1$, in the middle $n=2$ and in the right $n=3$. The volume of the 3D models is bounded by the xy-plane. When $n=2$ the curve is part of an ellipse and when $n=1$ the curve is a straight line (Erikson, 2004).

In the study conducted by Hung et al. (2012) in Australia (Queensland), an ellipsoid (equation 2.2) has been used to approximate the shape of acacia (*Acacia nilotica*), parkinsonia (*Parkinsonia aculeata*) and mesquite (*Prosopis pallida*).

$$\frac{x^2}{r_1^2} + \frac{y^2}{r_2^2} + \frac{z^2}{r_3^2} = 1 \quad (2.2)$$

After simulating the interaction of the sun with the tree crown, the 3D model is then projected onto the xy-plane in the viewing direction of the sensor.

Radiometric aspects

Once the shape function is constructed, the interaction of the scene irradiance with the tree crown needs to be modelled, which depends on both the shape and the physical composition of the crown. The simplest model assumes that the crown is an opaque surface (Pollock, 1996). This is a reasonable assumption for broad-leaved trees with leaves concentrated near the crown envelope, forming a continuous surface. However, this assumption is generally difficult to justify, particularly for mature trees growing under competition. To simulate the radiance of a permeable volume (i.e. a volume of light-scattering leaves), prior knowledge of the leaf area distribution function is needed. For this reason, in what follows, an opaque volume will be assumed.

In the opaque surface model, the contribution of the radiance leaving a surface cell in a particular viewing direction can be described by the Minnaert reflectance function (equation 2.3) (Minnaert, 1941):

$$L = \rho E_0 \frac{k+1}{2} [\cos(\theta_i)]^k [\cos(\theta_e)]^{k-1} \quad (2.3)$$

E_0 is the perpendicular irradiance of a beam of light with the incident angle θ_i (Mather and Tso, 2016; Pollock, 1996). θ_i is the angle between the surface normal of the tree and the beam direction (incident angle) and θ_e angle between the surface normal and the viewing direction (exitant angle) (figure 2.5). Both θ_i and θ_e are restricted to $[0, \frac{\pi}{2}]$, because if θ_i is outside this interval, the beam does not irradiate the surface point and if θ_e is outside this interval the surface point is not visible by the sensor. ρ is a variable between 0 and 1 which accounts for light absorption and k is a parameter between 0 and 1 related to the intrinsic properties of the surface. If $k = 0$, the surface is assumed to be perfectly smooth and if $k = 1$, the expression becomes the Lambertian reflectance function, where the surface is assumed to be perfectly rough (Mather and Tso, 2016; Pollock, 1996).

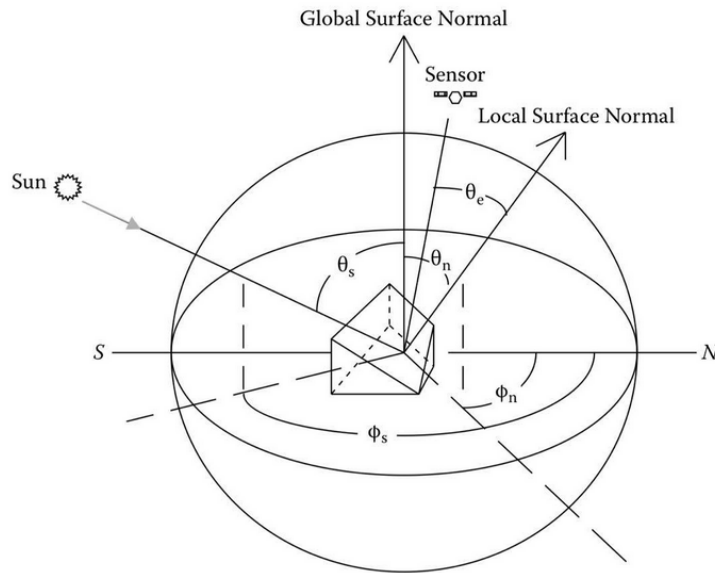


Figure 2.5: Geometrical relationships among the sun, the sensor, and the target position. Mather and Tso (2016)

Several studies (Gomes and Maillard, 2014; Gomes et al., 2018; Maillard and Gomes, 2016; Hung et al., 2012) have successfully generated tree crown models under the assumption of a perfectly rough crown. However, it is important to note that most surface materials have a roughness that lies between these two extremes and is wavelength dependent (Mather and Tso, 2016). Gomes and Maillard (2014) has introduced a simplified model (equation 2.4) in which the crown irradiance is calculated with the Lambertian reflectance model but has included an additional term accounting for diffuse ambient light.

$$L = L_{max} \cos(\theta_i) + amb \quad (2.4)$$

L_{max} is the maximum radiance of the tree and θ_i is the same as in formula 2.3. Amb is the ambient, diffuse light.

Larsen (1997) has extended the model introduced by Pollock (1996) by including the ground surface irradiance and the shadow cast by trees on this surface (figure 2.6 a and b). The ground plane is assumed to be a horizontal, flat surface that reflects and absorbs light. The light interaction can also be described by equation 2.3, but with ρ and k characterising the ground surface. Larsen and Rudemo (1998) have suggested that the surface can be approximated as a Lambertian surface (i.e. $k = 1$). The tree's shadow can be generated using ray-tracing with a simple hard shadow model (Gomes et al., 2018). The incident light vectors that are tangential to the surface function are projected to the ground plane to produce the outline of the shading (Hung et al., 2012). Maillard and Gomes (2016) has stated that using the whole shadow is not always beneficial in situations where the shadow is (partially) projected onto another tree. For this reason, a shadow clipping factor has been introduced (figure 2.6 c and d). The shadow cast by the tree on the ground can provide valuable information, which can be exploited by the template matching algorithm (Hung et al., 2012).

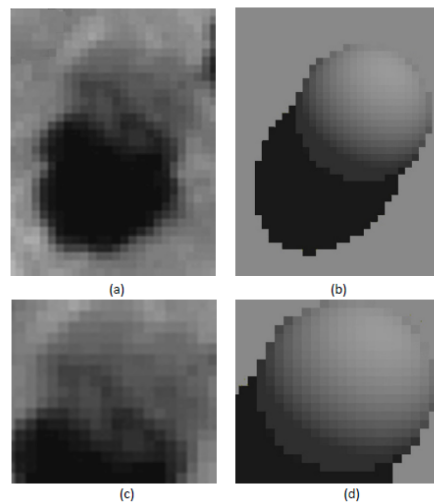


Figure 2.6: Comparison between an isolated tree from the (a) WorldView-2 image and (b) the geometrical-optical 3D model. A clipping factor of about 80 % was applied to the same images in (c) and (d) to enable the use of only a portion of the shadow in cases where that shadow is not projected on the ground but on another object (Gomes and Maillard, 2014).

2.3.2 Similarity measure

The templates generated in the first step are then used to search for the locus of best matching where trees are most likely to be, using a similarity measure and a threshold (Wang et al., 2004; Cheng and Han, 2016). Commonly used measures are the sum of absolute differences (SAD), the sum of squared differences (SSD), normalized cross-correlation (NCC) and Euclidean distance (ED) (Cheng and Han, 2016; Hisham et al., 2015; Fouda, 2014). SAD and SSD are computationally fast but are sensitive to outliers and not robust to template variations. On the other hand, NCC is computationally slower but is more accurate and robust under uniform illumination changes (Hisham et al., 2015; Fouda, 2014). Moreover, NCC has shown promising results in previous remote sensing studies using VHR

2. Literature review

and TM (Larsen et al., 2011; Gomes and Maillard, 2013; Shen and Bao, 2014; Korpela et al., 2004).

$$NCC(x, y) = \frac{\sum_i \sum_j [w_{(i,j)} - \bar{w}] \sum_i \sum_j [f_{(i,j)} - \bar{f}]}{\{\sum_i \sum_j [w_{(i,j)} - \bar{w}]^2 \sum_i \sum_j [f_{(i,j)} - \bar{f}]^2\}^{1/2}} \quad (2.5)$$

The NCC ranges from -1 to 1 and can be calculated using equation 2.5, where $w_{i,j}$ is the intensity of the pixel, \bar{w} is the mean intensity for the pixels under the template, $f_{i,j}$ is the intensity for the corresponding pixel in the template and \bar{f} is the mean intensity of the template (Vahidi et al., 2018; Erikson and Olofsson, 2005). Next, the template is slid across the image (figure 2.7), resulting in a correlation matrix. A higher value indicates a better match between the template and the image at that particular position. The NCC values were thresholded to find the centre of the tree crowns. Hung et al. (2012); Vahidi et al. (2018) have stated that the threshold should be selected with care because setting the threshold too low will lead to too many falsely detected trees while setting it too high leads to many undetected trees.

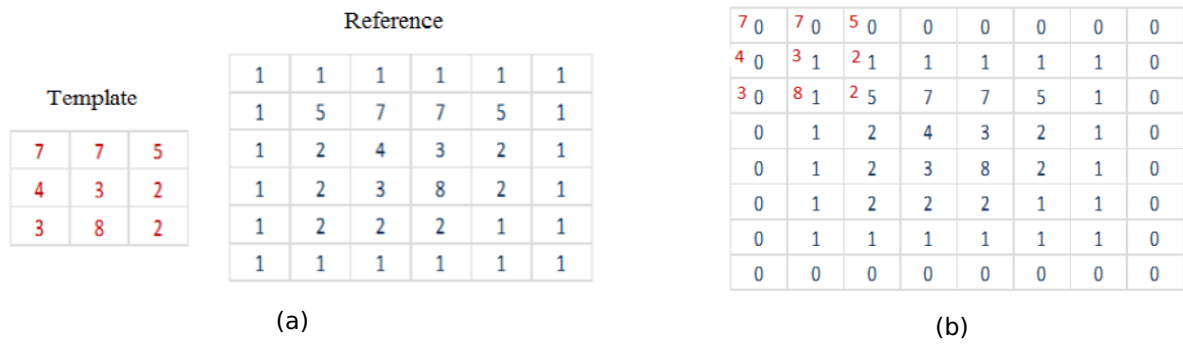


Figure 2.7: Figure a shows example matrices of a template and reference image. In figure b, the template is matched to the image. To do so, the reference image needs to be padded with zeros, so the centre of the template is placed on the first pixel of the reference image. At each position of the padded image, the NCC is calculated (Hisham et al., 2015).

2.4 Local maximum filtering

Local maximum filtering is a technique for identifying tree crowns in high spatial resolution imagery, which assumes that points with the highest brightness (i.e. digital number) within a search window are possible tree locations (Pouliot et al., 2002; Quackenbush et al., 2000). This method is adequate for trees with the greatest reflectance at their top, surrounded by lower-intensity pixels (Maillard and Gomes, 2016). For this reason, it has been widely used for detecting conifers. The windows that move over the image can be overlapping or not; the non-overlapping window ensures that pixels in the image are evaluated in only one sample window frame, reducing detection errors caused by portions of bright trees being repeatedly detected (Pouliot et al., 2002).

The success of the LM tree recognition depends on the careful selection of the search window. If it is too small, errors of commission occur by selecting nonexistent trees or

multiple radiance peaks for an individual tree crown. On the contrary, if the window size is too large, the LM filter will miss some trees (Pouliot et al., 2002; Wulder et al., 2000). When using a fixed window size, Pouliot et al. (2002) has suggested that the size should be smaller than the average crown size to reduce commission errors. To minimise this problem, Wulder et al. (2000) has introduced windows with varying sizes by assessing the spatial structure of the image obtained by analysing the local semivariogram with different pixel lags. By incorporating spatial information into the determination of window sizes, the commission errors (i.e., detections of trees that do not represent trees) will decrease, while omission errors are primarily in function of the image resolution (Wulder et al., 2000). A remotely sensed image can be processed for semivariance through the computation of relationships between pixels pairs (Curran and Atkinson, 1998). The larger the semivariance, the less similar the pixels are. Consider a transect passing across a remotely sensed image where the digital number z of pixels x is extracted at regular intervals ($x = 1, 2, \dots, n$) (Curran and Atkinson, 1998). The relationship between a pair of pixels found h pixels apart (the lag distance) is recorded as the average squared difference between all pixel pairs at lag distance h (Wulder et al., 2000). Within each transect, m pairs of observation are separated by the same h (Lévesque and King, 1999). For a given lag h , the semivariance is calculated by equation 2.6:

$$\gamma(h) = \frac{1}{2m} \sum_{i=1}^m [z(x_i) - z(x_i + h)]^2 \quad (2.6)$$

The semivariogram is a graphical representation of the semivariance as a function of lag (Lévesque and King, 1999) (figure 2.8). The semivariance will rise until the 'sill', which indicates the maximum variability between pixels. The range is the distance to the sill (Curran and Atkinson, 1998). The pixel values at the lag locations greater than the range considered spatially independent of the values within the range (Wulder et al., 2000). According to Woodcock et al. (1988), the range is often associated with the size of individual objects, such as tree crown size in VHR imagery or forest stand size in low-resolution imagery. Wulder et al. (2000) has suggested the computation from transects in the eight cardinal directions around each pixel in the image. For each of the eight transects, the semivariance at each lag is stored in an array until the semivariance values cease to increase, representing the range in that direction. The eight directional range values are then averaged, resulting in the range for that pixel location. Once the semivariance ranges are obtained, a conversion into window sizes is needed (Daley et al., 1998) using a conversion table. Wulder et al. (2000); Pouliot et al. (2002) have found that by calculating a personalised window size for each pixel, greater accuracy can be obtained compared to using a fixed window size is used when counting trees in coniferous forests.

A second problem with the LM filter is that it is affected by false bright pixels (Maillard and Gomes, 2016). Brandtberg and Warner (2006); Pouliot et al. (2002) have suggested applying a Gaussian filter to the image to lower the number of local maxima being identified. This low-pass filter grants more weight to the crown centre pixels than those located toward the crown edge, which might belong to other bright objects or noise (Maillard and Gomes, 2016). The smoothing scale directly affects the number of local maxima and also causes a rounding-off of the brightness values of the tree crown edges (Brandtberg and Warner, 2006). Pouliot and King (2005) have adapted a method first proposed by Culvenor (2000) to

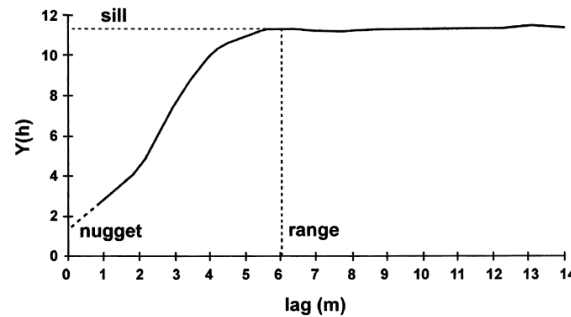


Figure 2.8: Example of a semivariogram where the semivariance (y) is plotted in function of the lag distance. The sill, the range and the nugget are indicated on the graph. (Lévesque and King, 1999)

determine the optimal smoothing factor by plotting the number of local maxima in function of the smoothing factor (σ) and finding the point where the function starts to diverge. The optimal global scale is defined as that which minimises both errors of omission and errors of commission, producing the best overall detection results (Pouliot and King, 2005). Pouliot and King (2005) have found by comparing different optimisation approaches for LMF (global optimum smoothing factor, variable local window size and local smoothing factor), the use of an optimal global smoothing factor produced the most consistent accuracies and is the least parameter dependent. A second approach is the use of the Laplacian-of-Gaussian (LoG) (Wang et al., 2004), which, because it is based on the second-order spatial derivatives, highlights locations where the image intensity varies rapidly (Brandtberg and Warner, 2006).

Wang et al. (2004) developed a more complex method for detecting tree tops in a coniferous forest in British Colombia. In the first step, the trees were separated from the background using a Laplacian of the Gaussian edge detector on the first principal component of principal component analysis. Each closed contour derived by edge detection is treated as an object containing one or multiple trees. A local non-maximum suppression method, which sets non-local maxima to zero on a grey-level image, was used to obtain one set of treetops. The second set of treetops was obtained by applying a local maximum filter on the morphologically transformed distance. This second set is based on the assumption that treetops are located near the tree's centre. In order to satisfy the two assumptions associated with tree-tops, the final markers were obtained by intersecting both. If a maximum-distance treetop is located within a three-by-three window surrounding a grey-level treetop, then this treetop is selected as a final treetop. Khan and Gupta (2018) have suggested a method suited for classifying trees in a dense forest in South India. The technique uses two morphological operations: first, opening and closing are applied to the image, followed by opening and closing by reconstruction. This process refines the image by removing unusual noise on which an LMF is applied.

3. MATERIAL AND METHODS

The concepts and two methods (template-matching and local maximum filtering) described in section 2 were adjusted and applied to the study area to detect *Melia Volkensii* trees in the farms and in the plantation. Firstly, in section 3.1 this area was described. Secondly, the satellite image and ground truth data collection were explained in section 3.2. Thirdly, the regions of interest (ROI) extraction before the application of the two approaches were discussed. Fourthly, in sections, 3.4 and 3.5, the template-matching and local maximum filtering approaches were elaborated. Lastly, the strategy used for the accuracy assessment of the methods was discussed in section 3.6.

3.1 Area description

The study focused on a 9 x 13 km area (figure 3.1) located in Kitui county in semi-arid Kenya. The plantation borders Lake Kiambere on the western side and the farms are located east of the plantation. In Kenya, rain is delivered during two seasons. One during the boreal spring (March-May), known as "the long rains" and one in the boreal fall (October-December), "the short rains" (Cook and Vizzy, 2013). Since 2006 daily rainfall data has been collected at the plantation site (Vandenabeele, 2021). Figure 3.2 shows the average monthly precipitation from 2006 until 2021, with the monthly precipitation in 2021 indicated since this is the year where most data were acquired. The seasonality throughout 2021 is shown in figure 3.3, where distinct red areas represent areas with a lot of vegetation.

3. Material and methods

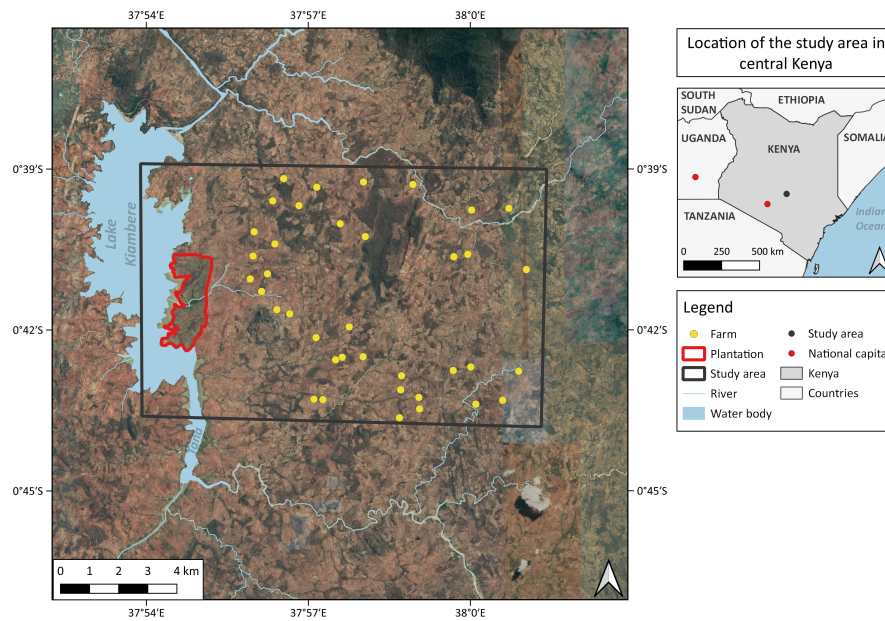


Figure 3.1: Location of the study area in central Kenya, with an indication of the plantation (red line), bordering on Lake Kiambere, and the 38 farms (yellow dots) (Forceville, 2022).

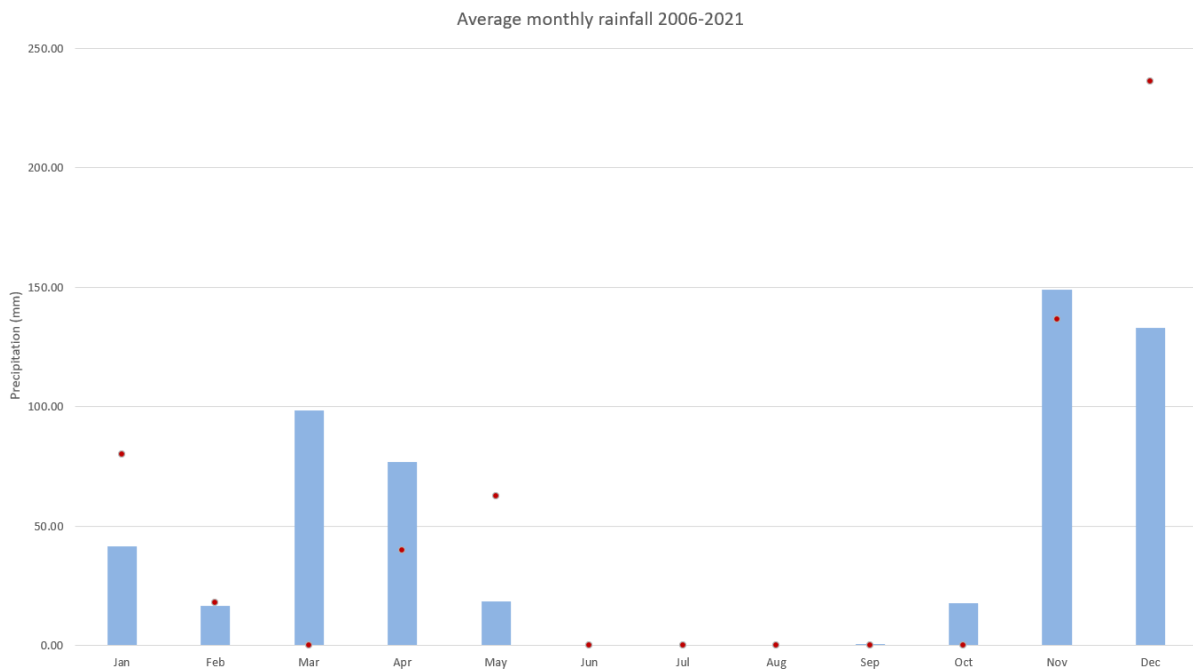


Figure 3.2: Average monthly rainfall from 2006 - 2021 in Kiambere. Monthly precipitation from 2021 is indicated by red dots (Vandenabeele, 2021).

3. Material and methods

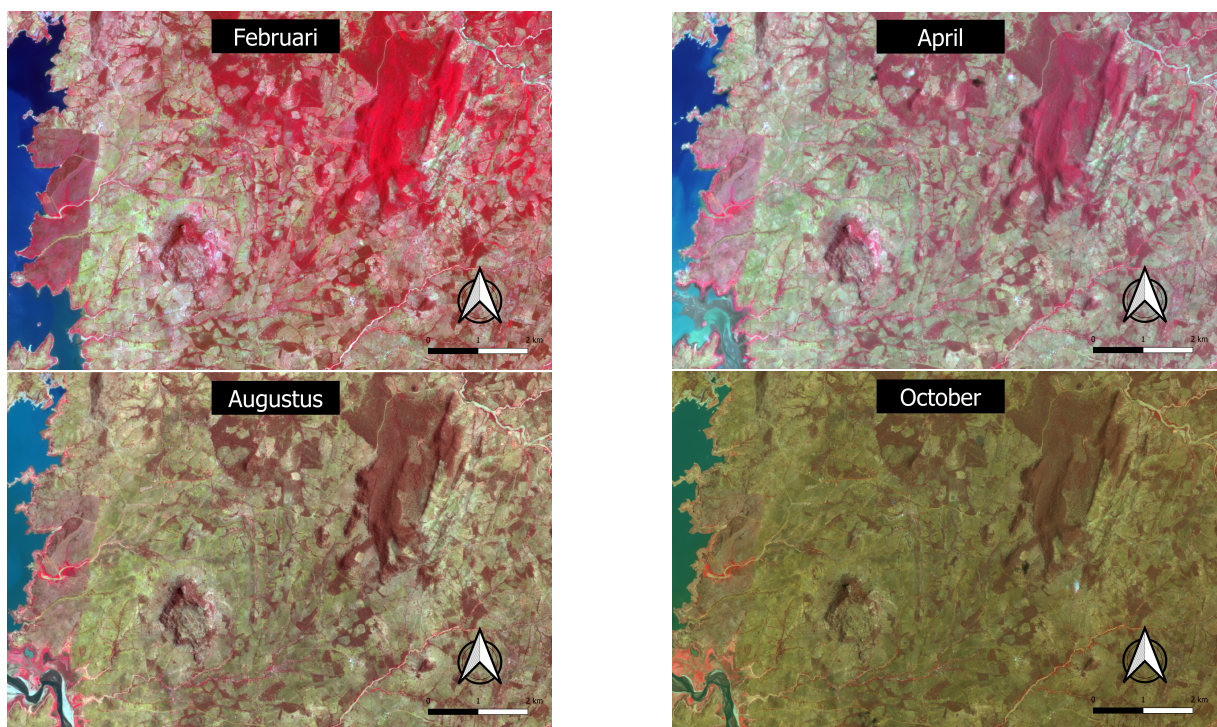


Figure 3.3: False colour composite with a NIR, red and green scheme causing vegetation to give the image a distinct red colour. The PlanetScope satellite acquired the imagery in February, April, August and October of 2021. The four false colour images display the seasonality of the study area throughout the year.

3.2 Data collection

3.2.1 Satellite imagery

The image data was acquired on the second of June 2021 at 7:51 AM local time (GMT +3) with the Pléiades 1B satellite. This specific satellite was chosen because of budgetary limits. The image contains one panchromatic band with a spatial resolution of 0.5 meters and four multispectral bands (Red (R), Green (G), Blue (B) and Near Infrared (NIR)) with a spatial resolution of 2 meters (table 3.1). The Pléiades 1B satellite is a sun-synchronous satellite with a swath width of 20km. At the time of acquisition, the solar azimuth varied from 43.911°-43.539° and the solar elevation varied from 57.430° - 57.546°. The imagery has a pre-processing level 1C, where the digital numbers were converted to a 16-bit image containing top-of-atmosphere reflectance. The ortho geometric processing level is ortho, which produced a georeferenced image in Earth geometry, corrected from acquisition and terrain off-nadir effects (Airbus, 2011). The imagery was registered to WGS 84 (Worldwide Geodetic System 84) datum and the Universal Transverse Mercator (UTM) zone 37S projection.

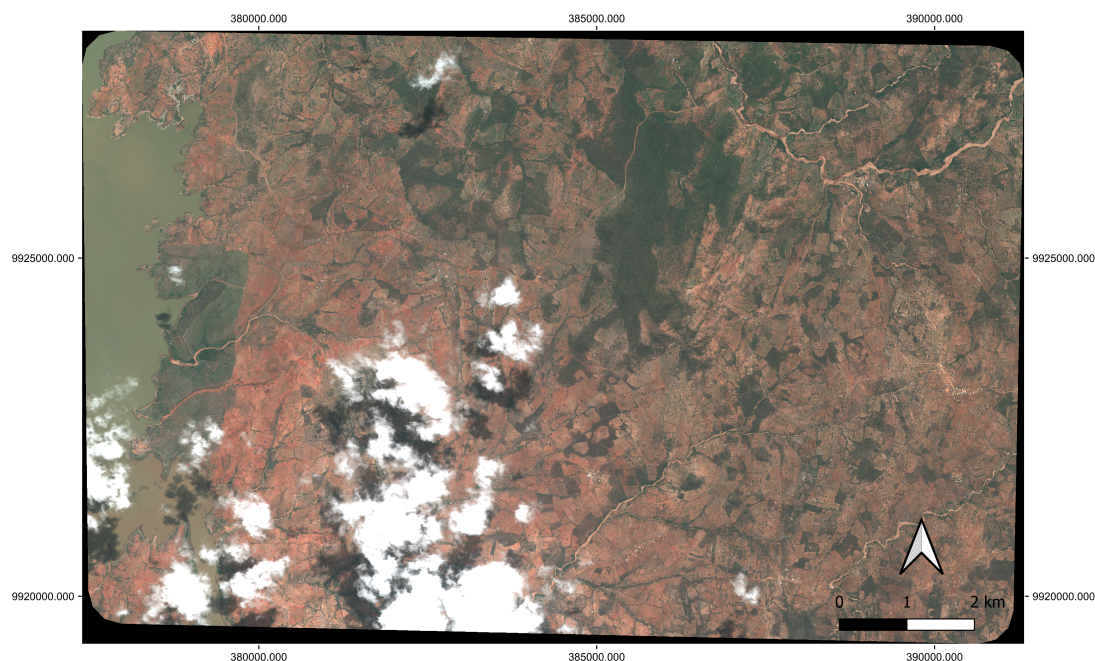


Figure 3.4: True colour image acquired by the Pléiades 1B satellite on the second of June 2021. The image extent is displayed in CRS WGS 84/UTM zone 37s and is approximately 9 km x 13 km.

Of all the measured *Melia* trees in the farms and the plantation, 99.22% have a diameter at breast height ranging from 0.00 cm to 25.00 cm (table 3.3), corresponding with an average crown diameter range of 131.57 cm to 745.11 cm. This corresponds with a crown diameter-to-pixel ratio between 0.66:1 and 3.73:1 in the multi-spectral bands. To obtain a ratio between 3:1 and 19:1 (section 2.1), a 0.50 m resolution pan-sharpened Pléiades image was produced. The pan-sharpened image was attained by fusing the 2 m resolution MS image with the 0.50 m resolution panchromatic image with the use of the Gramm-Schmidt

3. Material and methods

Table 3.1: Band parameters of Pléiades 1B satellite.

Band	Bandwidth (nm)	Spatial resolution (m)
Panchromatic	480-830	0.5
Blue	430-550	2
Green	490-610	2
Red	600-720	2
NIR	750-950	2

spectral sharpening method in the SPEAR pan-sharpening tool in ENVI 5.6.1. The sharpened imagery resulted in a ratio between 2.63:1 and 14.90:1, which was still unsatisfactory for the smaller trees.

3.2.2 Field survey

Ground truth data were collected on the plantation and on farms during a field campaign from October 2021 - December 2021 and during a second, shorter one in January 2022. The plantation was divided into grid cells of 10 ha. In each grid cell, a sampling plot was randomly located and in grid cells with a tree coverage larger than 5 ha, an additional plot was selected (table A.1). In total, 64 square inventory plots were measured on the plantation. Farms with less than 30 seedlings were not visited; consequently, only 17 out of the 34 farms were visited. On these 17 farms (table A.2), a full inventory of the *Melia volkensii* trees belonging to BGF was conducted. In both the farms and the plantation, locations covered by clouds on the Pléiades satellite image were not visited. The following tree variables were measured or taken note of in both the plantation and the farms:

- *The location* was measured using a garmin 60scx GPS during the first field campaign. During the second field campaign, the location was measured with a phone. The coordinates were notated in CRS WGS84/ UTM zone 37S. In five plots in the plantation, light detection and ranging (LiDAR) scans were taken with an iPad Pro and later, locational data was extracted from these scans.
- *The diameter at breast height (cm)* was measured using a measuring tape.
- *The crown diameters (m)* were measured in N-S and E-W directions using a measuring tape.
- *The height (m)* was measured using the phone application 'Arboreal - Tree height' during the first field campaign and with a TruPulse dendroscope during the second field campaign.
- *The tree age (years)* was taken note of. In the plantation, this is indicated on the tree trunk. On the farms, it was noted when the trees were planted.
- *The species name* was taken note of based on expert knowledge.
- *The soil degradation* due to water erosion was recorded based on visual assessment. The tree was given a score depending on the type of soil loss, namely no erosion - 0, sheet erosion - 1, rill erosion - 2 and gully erosion - 3 (Omuto, 2008).

3. Material and methods

- The tree vitality was given a score from 1 (good) to 5 (dead) based on expert knowledge.

The species frequency of the total amount of collected data per afforestation layout can be found in table 3.2. In table 3.3, the descriptive statistics of *Melia volkensii*. The trees were divided into diameter classes of 5 cm based on the diameter at breast height (DBH). Per class, the table contains the average value of the crown diameter, the height, the tree age and the tree vitality. The last column shows the percentage of trees per diameter compared to the total amount of *Melia Volkensii*. In the plantation, one outlier can be noticed; this could be a measuring error. The same descriptive statistics of Neem, Acacia and Baobab are included in the appendix (table B.1) since these are less important for the remainder of this study.

Table 3.2: Species frequency (%) of the collected data for each afforestation layout.

Species		Frequency in field data (%)	
Scientific name	Common name	Plantation	Farms
<i>Melia volkensii</i>	Mukau	78.91	100.00
<i>Azadirachta Indica</i>	Neem	17.64	0.00
<i>Acacia sp.</i>	Acacia	3.32	0.00
<i>Adansonia sp.</i>	Baobab	0.12	0.00

Table 3.3: Descriptive statistics of *Melia Volkensii*. The trees were divided into diameter classes of 5 cm based on the diameter at breast height (DBH). Per class, the table contains the average value of the crown diameter, the height, the tree age and the tree vitality. The last column shows the percentage of trees per diameter compared to the total amount of *Melia Volkensii*.

DBH (cm)	Crown diameter (cm)	Height (cm)	Tree age (year)	Tree Vitality	Number of trees (%)
Plantation					
0.0 - 5.0	131.57	3.87	6.12	1.97	1.78
5.0 - 10.0	262.00	7.33	8.34	1.65	14.59
10.0 - 15.0	410.98	9.73	9.43	1.28	46.98
15.0 - 20.0	521.19	11.03	10.15	1.12	30.76
20.0 - 25.0	625.71	12.31	10.61	1.18	4.99
25.0 - 30.0	644.64	10.83	13.82	2.09	0.58
30.0 - 35.0	676.70	10.54	13.80	1.2	0.26
35.0 - 40.0	NA	NA	NA	NA	0.00
40.0 - 45.0	NA	NA	NA	NA	0.00
45.0 - 50.0	1150.00	12.20	15.00	1.00	0.05
Farms					
0.0 - 5.0	163.43	12.2	2.85	1.71	16.82
5.0 - 10.0	300.70	5.35	3.36	1.57	45.60
10.0 - 15.0	441.33	7.63	3.96	1.28	28.25
15.0 - 20.0	590.69	9.11	4.78	1.20	6.83
20.0 - 25.0	745.11	9.80	5.00	1.36	1.84
25.0 - 30.0	718.90	12.20	5.00	3.00	0.66

3.3 Regions of interest

3.3.1 Cloud mask

The cloud-covered areas and areas lying in the shade of the clouds on the Pléiades image were excluded from executing the ITCD algorithms. Airbus provided a cloud mask with the image delivery.

3.3.2 Tree mask

Before the detection of crowns, extraction of the forested areas is often performed (Bunting and Lucas, 2006). Non-forested areas such as soil, shrubs, grass, clouds or other material are frequently represented as bright pixels causing falsely detected trees (Brandtberg and Walter, 1998). However, the plantation mostly consists of a forested area with little to no bare soil. For this reason, the area is easily manually delineated and a shape file from BGF was already available for the plantation in Kiambere. The delineation of trees in the farms almost comes down to the individual delineation of tree crowns since the planting distance is large to allow intercropping with crops. For this reason, the farms are also manually delineated even though they still contain a lot of non-forested areas. Furthermore, this study intends not to count all trees on the image, but only the trees on certain farms, so it also serves as a region of interest.

Nevertheless, a tree mask was still created but used for the accuracy assessment as described in section 3.6. The classification of non-forested and forested areas was the focus of the second objective of this project. Forceville (2022) followed a multi-step procedure by Vahidi et al. (2018) to mask out the candidate tree crown (CTC) class objects from non-tree crown (NTC) class objects. Some adjustments were made to the original procedure due to the lack of height information. In the first step, object primitives, which do not yet have any meaning, were created in eCognition® Developer version 10.2 using a multiresolution segmentation algorithm. In the second step, these object primitives were classified into CTC or NTC. The elevated vegetation (rough texture) was separated from the understory (smooth texture) by performing an edge extraction Lee Sigma filtering for detecting bright and dark edges in the R band (sigma value = 5). A new band (LeeSigmaSum) was created by adding the bright edge Lee Sigma band to the dark edge Lee Sigma band. The LeeSigmaSum was then filtered with a Gaussian smoothing filter with a kernel size of 25 x 25 pixels, resulting in the Roughness band. Lastly, the CTC class was separated from the NTC by thresholding the R to Roughness ratio and the normalised vegetation index (NDVI, equation 3.1) (Vahidi et al., 2018). For more detailed information about the procedure, turn to Forceville (2022).

$$NDVI = \frac{NIR - R}{NIR + R} \quad (3.1)$$

3.4 Template matching

The main algorithm is the template matching algorithm (algorithm 1), which was programmed in Matlab[®] R2022A (as are the following). This algorithm locates the centre of trees on a greyscale image within certain regions of interest (ROIs) using k amount of templates. In the first step, the image is cropped using the ROIs, which will decrease the computation time substantially. The following steps are thus repeated q times for each ROI separately. In the second step, the similarity between the image and each template was calculated using normalised cross-correlation (formula 2.5 section 2.3.2). In the third step, the correlation matrix of each template is scanned to spot where it exceeds the threshold similarity value. If the latter is the case, that point is a hit and marked as 1. In the fourth step, multiple hits that probably originate from the same tree (false positives) are removed using a minimum distance (figure 3.5), expressed in pixels (Erikson and Olofsson, 2005). When two or more hits are too close to each other, the templates will overlap. The lower the minimum distance, the more overlap between templates is allowed. This distance can be user-defined based on expert knowledge, or the default distance can be used. The default allows templates to have a horizontal or vertical overlap of one pixel. When two or more templates exceed the maximum overlap, the tree with the highest similarity to the template is taken and all others are removed. The default minimum distance was selected in the farms, while in the plantation, a minimum distance of 5 pixels (2.5 m) was chosen as there is a fixed planting distance of 4 m in the plantation. In the final step, the points of each ROI are merged together. The result is thus a georeferenced logical matrix where one represents the centre of a tree.

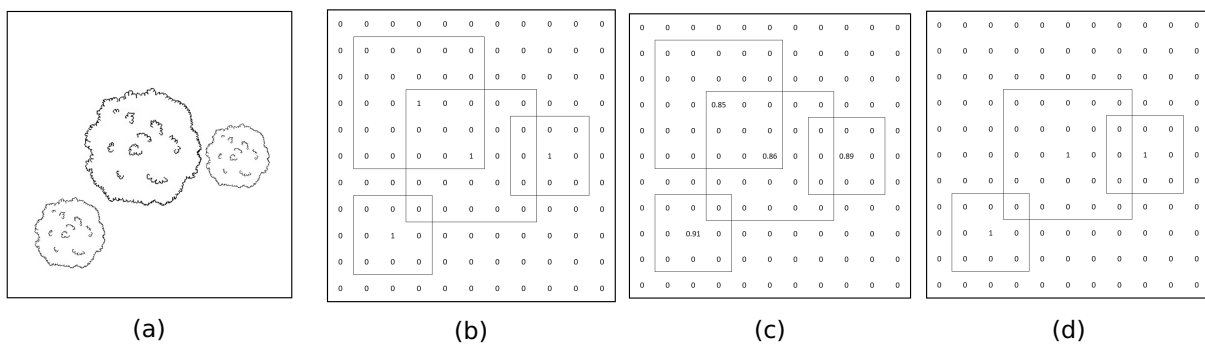


Figure 3.5: Example of how the default minimum distance works. Figure b shows the positions (=1) with a correlation value higher than the threshold. The top left and middle hit lie within the default minimum distance and detect the same tree (a). The middle hit shows a higher correlation (c), so the top left one is deleted. Figure d shows the corrected matrix.

In sections, 3.6 and 3.4.1 the parameter optimisation and accuracy assessment will be further elaborated.

Algorithm 1: Template Matching Algorithm

Data: greyscale image (I), geodata (R), k templates (T), threshold (Thr), minimum distance (d_{min} , optional), q regions

Result: Tree locations in selected regions

for $q = 1: \text{regions}$ **do**

 Crop I using the selected region.;

for $i = 1:k$ **do**

 correlation = normxcorr2(T_i, I) ; // calculate normalized cross-correlation between I
 and T_i

end

for $i = \text{size}(I,1)$ **do**

for $j = \text{size}(I,2)$ **do**

if correlation > Thr **then**

 points(i,j) = 1 ;

// tree at location

else

 points(i,j) = 0 ;

// no tree at location

end

end

end

for the entire points matrix do

 Eliminate points lying within the minimum distance of each other.

if closer than the minimum distance then

 Select the tree with the highest correlation

end

end

 The corrected points matrix can be located back onto the original image using geodata (R).

end

Tree locations = assembly of all georeferenced points matrices of each region.

3.4.1 Parameter optimisation

In the process described above, there are several critical inputs, namely the input greyscale image, the templates and the threshold. The choice of these inputs largely influences the performance of the template matching algorithm; thus, these parameters were optimised or tested. Figure 3.6 shows the optimisation process, which will be described in the following paragraphs.

3. Material and methods

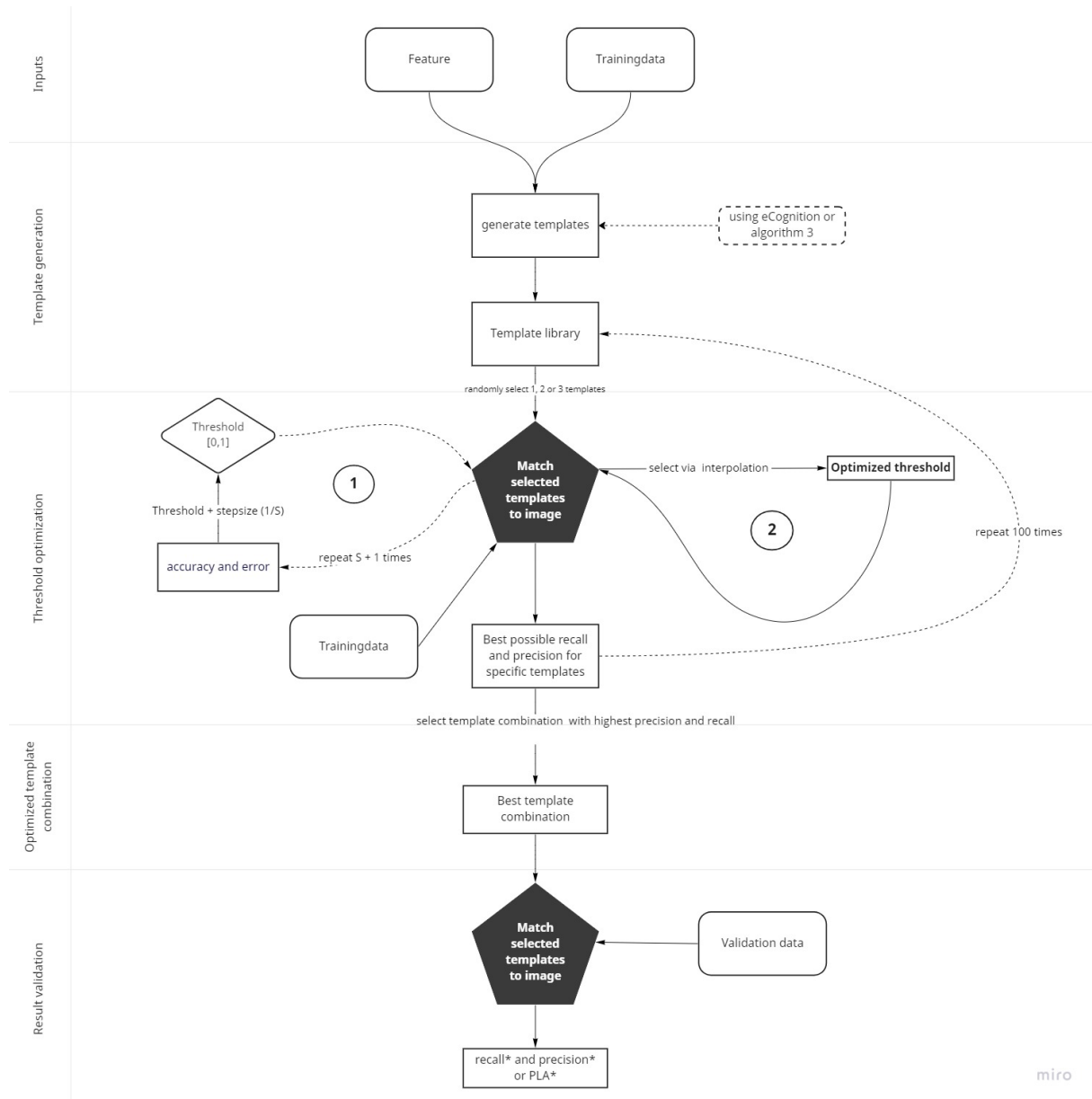


Figure 3.6: Workflow of the optimisation of the template matching algorithm. Firstly, a set of greyscale images (or features) will be selected. For each feature, two sets of template libraries will be generated using algorithm three or eCognition[®]. Secondly, for that set feature, the template combination resulting in the highest precision and recall is obtained by randomly selecting three templates and simultaneously optimising the threshold. Lastly, the optimised algorithm is validated using a validation data set.

The image

The first critical input is the greyscale image. Forceville (2022) has ranked 68 spectral, textural and shape object features from most suitable to least suitable for the distinction between candidate tree crowns (CTC) and non-tree crowns (NTC) for this given area. In figure 3.7, the density functions of the classified objects (i.e. CTC and NTC) for the top five highest-ranked features are plotted, visualising the separability between the two classes. These five features are image object statistics based on a single band or a combination of spectral bands (Benz et al., 2004). The mean band value (mean G, mean R, Mean B) represents the mean brightness of an image object within a single band (Mishra and Crews, 2014). The standard deviation (StDev NIR) is the standard deviation of all pixels forming an image object within a band (Mishra and Crews, 2014). The brightness represents the mean value of the spectral mean values of the R, G, B and NIR bands of an image object (Van Coillie et al., 2007). For this study, the top five features were tested to see which results in the highest performance of the TM algorithm. However, because this is not an object-based procedure, the features were calculated slightly differently: the true values of the R, G, B and NIR bands were used instead of the mean or standard deviation and the brightness was calculated as the mean value of the four spectral bands (i.e. B, G, R and NIR) of each pixel.

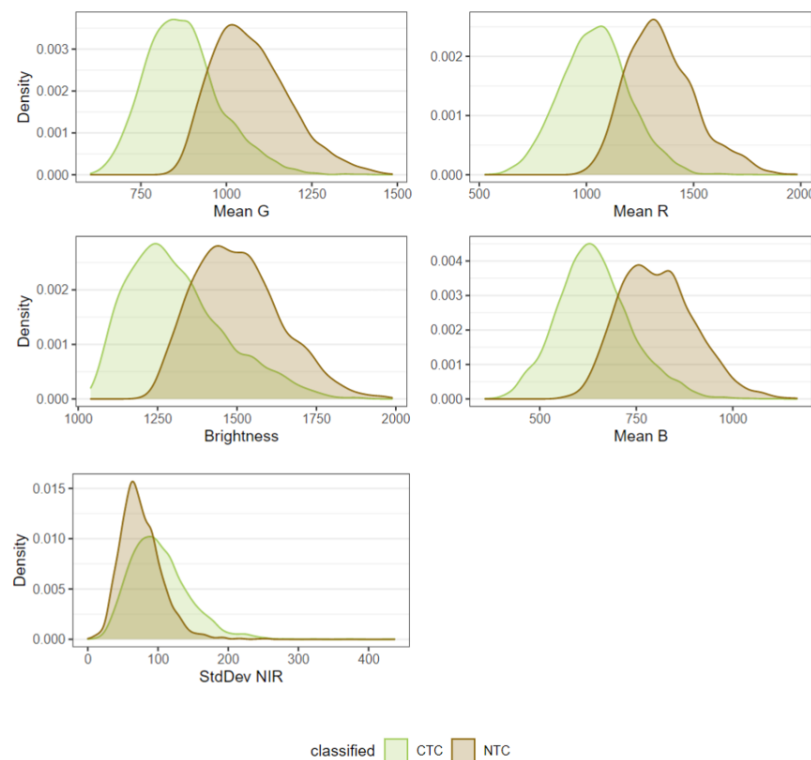


Figure 3.7: Density functions of the classified CTC as well as NTC objects of the five highest-ranked features (Forceville, 2022).

Threshold

The second critical input is the threshold. Lower threshold settings lower the decision boundaries and thus return more detections but increase the number of false positives (i.e. decrease in precision) (Hung et al., 2012). Contrarily, higher settings reduce the number of detections and consequently increase the number of false negatives (i.e. lower true positive rate). The threshold value was optimised for a given greyscale image and a given set of templates using algorithm 2. Firstly, this algorithm calculates the true positive and false discovery rate S times for each threshold in the $[0, 1]$ range with a predefined step size, which is the inverse of the number of iterations (S). The recall and precision were calculated using the training data. In this study, the number of iterations was set to 10. Secondly, an exponential function was fitted through the recall and the precision. Both these functions needed to be maximised in order to find the optimal threshold. Lastly, the intersection of these two functions is considered the optimal threshold.

For the threshold optimisation in the plantation, solely the LiDAR data was used since this is the only positional accurate data available.

Algorithm 2: Threshold Optimization

Data: greyscale image (I), geodata (R), k templates (T), number of iterations (S), minimum distance (optional), regions, validation data, tree mask

Result: Optimal Threshold

```
Thr = 0 ; // Initialize threshold starting at zero
for i = 1:S+1 do
    [recall(i), precision(i)] = TMacc(I,R,T1, ... Tk, Thr, regions, validation data, tree mask); // Calculate precision and recall using the TMacc function.
    Thr = Thr + 1/S ; // Increase Thr with the stepsize (=1/S)
end
```

Fit an exponential function through the recall (f_1) and precision (f_2) array;
Obtain the *Optimal Threshold* by finding the threshold where both functions intersect.;

Templates

Since it was not possible to use a large template library because this increased the computation time significantly, it was thus favourable to find an optimal combination of templates, resulting in the highest performance of the TM algorithm. A maximum of 3 templates were selected simultaneously because a higher number increased the computation time. Moreover, trial and error have shown that increasing the number of subgroups above three did not increase the correlation between the generated templates and the training data further. Firstly, two sets of templates were generated, empirical and synthetic, as will be explained in the next two paragraphs. Secondly, the best combination of these two sets was found by randomly selecting 100 combinations of one to three templates. For each combination, the threshold was first optimised as described in section 3.4.1.

The empirical templates were created with eCognition® version 10.2 using the training data. eCognition® allows for the choice in window size and group size. If the template type 'grouped' is selected, the algorithm attempts to find subgroups in the training data and creates a separate template for each subgroup (Trimble, 2022). Templates with the following window size were created: 9, 13, 15, 17 and 21. For each of these window sizes,

3. Material and methods

varying group sizes (1-3) were selected, resulting in 30 individual templates. In addition to the random selection of templates, 'the standard' combination was also tested. This is the combination of templates as they were generated in eCognition®.

A shape function and a radiometric model were needed to generate the second set of templates, as described in section 2.3.1. The shape function was approximated by formula 2.2, which assumes that the tree crowns are rotationally symmetric about the vertical axis (i.e., equal r_1 and r_2). To simulate the interaction of the tree crown with the scene irradiance, the model (equation 2.4) introduced by Gomes and Maillard (2014) was chosen, as described in section 2.3.1. The tree's shadow was generated using the simple hard shadow model (Hung et al., 2012; Gomes et al., 2018). The clipping factor was implemented with the buffer variable, so only part of the shadow was used. The clipping factor in the agroforestry was set to 4 pixels and to 0 in the plantation. This creates a background buffer of 4 and 0 pixels around the tree crowns. The tree crown models were projected onto the xy-plane, assuming the sensor was directly above the simulated tree.

The shape function and the radiometric model were combined in algorithm 3. The inputs of this synthetic template generation algorithm are the following:

- *the pixel size (m)*
- L_{max} : the maximum brightness of a tree. This was estimated by calculating the mean maximum brightness of the training and validation data.
- *amb*: diffuse light was estimated by sampling 50 points in cloud-covered areas of the image and taking the mean of these values (Gomes et al., 2018).
- *BR*: the mean background reflectance was estimated by sampling 50 random points on the image and taking the mean of these values.
- θ : the sun elevation (degrees)
- ϕ : the sun azimuth (degrees)
- D : the mean crown diameter (m)
- h_1 : the total tree height (m)
- h_2 : the crown height (m)
- *buffer*: a variable determining how many additional background pixels will be added to the crown diameter.

The input tree parameters were estimated by finding the relation between the mean crown diameter, crown height and the total tree height of the data and the diameter at breast height. For each upper limit of the diameter classes (table 3.3), the corresponding tree parameters were estimated and a template was generated.

3. Material and methods

Algorithm 3: Synthetic template generation

Data: L_{max} , amb, BR, the pixel size (m), θ , ϕ , D , h_1 , h_2 , buffer

Result: Template

Generate an ellipsoid with dimensions D and h_2 and h_1 above the xy-plane.

for each facet on the ellipsoid do

θ_i = the angle between the local surface normal and the incident angle of the sun.

if $\theta_i > 90^\circ$ **then**

 | $L = \text{amb}$ // this facet of the tree is not directly illuminated

else

 | $L = L_{max} * \cos(\theta_i) + \text{amb}$

end

end

Generate the shadow of the tree crown by finding the incident light vectors that are tangential to the ellipsoid and projecting them onto the ground plane.

Project the tree crown and the shadow onto the xy-plane with the pixel size using linear interpolation.

Give all background pixels the background radiation BR.

Crop the generated template so it has a window size of the maximum diameter + the buffer. If this results in an even window size the 1 row and 1 column are added.

3.5 Local maximum filtering

Since simply applying a local maximum filter (LMF) on the image would lead to many falsely detected trees, a pre-processing method was tested and evaluated. The method was tested on the same five features as in the TM method. The pre-processing method is an adjusted version of the procedure proposed by Khan and Gupta (2018) to detect trees in a dense forest. In several consecutive morphological reconstruction operations, the image is filtered and smoothed to enhance the treetops by removing small holes and dots that are usually noise. The morphology-based tree detection is implemented as follows:

1. Opening
2. Closing
3. Opening by reconstruction
4. Closing
5. Closing by reconstruction
6. Apply non-overlapping local maximum filter

A disk with a diameter of 5 pixels was used as a SE for morphological operations. The non-overlapping LMF was carried out using a moving 6 x 6 window.

3.6 Accuracy assessment

True positives (TP) were calculated as every tree from the validation data set that was found. False negatives (FN), were calculated as every validation point that was not found. Some problems surfaced when trying to determine the false positives (FP) and true negatives (TN). To calculate FP, a full inventory of the selected farms was needed, but not available. To partially solve this problem a tree mask, created by Forceville (2022), was used as described in section 3.3.2. Every positive point, which does not lie inside the tree mask is considered false. However, this approach assumes that the tree mask is 100 % accurate. Nevertheless, this was not the case and some mistakes were introduced. TN are the negative detections lying in the area between the trees (Hung et al., 2012). Since no negative points were generated by the TM algorithm nor the LMF, it was not possible to calculate TN. This problem was omitted by using precision and recall which does not require counts on TN (Hung et al., 2012). The tree detection algorithm performs well if the recall or true positive rate (equation 3.2) is high and the false discovery rate (equation 3.3) is low. The lower the FDR, the higher the precision (equation 3.4). The final performance (F-score) was determined by the harmonic mean of the two measures (equation 3.5) (Salamí et al., 2019). The F-score, also known as F_1 -score or F-measure, is a measure that combines precision and recall to evaluate the performance of a test or model Manning et al. (2008). The precision is the proportion of correctly identified positive outcomes among all positive predictions. The recall, also known as sensitivity or true positive rate, is the proportion of correctly identified positive outcomes among all actual positive outcomes. The F1 score is calculated as the harmonic mean of precision and recall, with a value of 1 indicating perfect performance and a value of 0 indicating the worst possible performance.

$$TPR = \frac{TP}{TP + FN} \quad (3.2)$$

$$FDR = \frac{FP}{TP + FP} \quad (3.3)$$

$$precision = 1 - FDR \quad (3.4)$$

$$F = \frac{2}{\frac{1}{precision} + \frac{1}{recall}} \quad (3.5)$$

For the accuracy assessment on the plantation, there is no tree mask available. The validation data also had a very low spatial accuracy of +/- 5m. For this reason, only a whole plot assessment was performed. In each plot a full inventory was available and the plots were manually delineated. However, this was not 100% accurately done, some mistakes were introduced. The plot-level accuracy (PLA) was computed by comparing the total number of trees (d) detected by the TM algorithm or LMF to the reference count (n) (equation 3.6). The closer the PLA is to 1, the better the algorithm performs. Lamar et al. (2005) have men-

3. Material and methods

tioned that the PLA can be misleading due to commission and omission errors cancelling each other out and providing no information about tree-location accuracy.

$$PLA = \frac{d}{n} \quad (3.6)$$

After the optimisation process as described in section 3.4.1, the algorithm was run again and validated with an independent validation dataset. This results in F-score* and PLA*.

4. RESULTS AND DISCUSSION

In the following sections (4.1 and 4.2), the results of the template-matching (TM) algorithm and the local maximum filter (LMF) will be discussed. The results of the TM algorithm in the different afforestation layouts are discussed separately, while those of the LMF are discussed together. Followed by a method comparison in section 4.3 and a recommendation segment in section 4.4.

4.1 Template matching

Firstly, the radiometric and geometric parameters estimation is needed for algorithm 3 as discussed in section 4.1.1. Secondly, the results of the template matching algorithm in the farms and the plantation are discussed separately in section 4.1.2 and 4.1.3.

4.1.1 Tree parameters

The input tree parameters as described in section 3.4.1 needed for the template generation algorithm 3 are firstly estimated. This comes down to estimating the following geometric parameters:

- D : the mean crown diameter (m)
- h_1 : the total tree height (m)
- h_2 : the crown height (m)

and the following radiometric parameters:

- L_{max} : the maximum brightness of a tree. This was estimated by calculating the mean maximum brightness of the training and validation data.
- amb : diffuse light was estimated by sampling 50 points in cloud-covered areas of the image and taking the mean of these values (Gomes et al., 2018).
- BR : the mean background reflectance was estimated by sampling 50 random points on the image and taking the mean of these values.

Geometric parameters

The input geometric parameters are estimated by finding the relation between the diameter at breast height (*dbh*) and the mean crown diameter and the *dbh* and the total tree height. The crown height shows the best relation with the total height instead of the *dbh*. The resulting empirical equations and their corresponding R^2 can be found in table 4.1. In both the farms and the plantation, the total height shows the best fit with the crown height with an R^2 of 0.83 and 0.81. The DBH in the farms shows a good fit with the crown diameter ($R^2 = 0.74$) and the total height ($R^2 = 0.74$), while in the plantation, this is much lower with an R^2 of respectively 0.59 and 0.41. The decreased fit could be due to the high crown closure and overlap in the plantation, which made it harder to accurately measure the total height and the crown diameter in the field. In the agroforestry layout, it was possible to distinguish tree crowns from one another since the trees were planted with wide spacing and hardly any crown overlap occurred.

Table 4.1: The empirical formulas used to estimate the crown diameter (D), the total height (h_1) and the crown height (h_2) of *Melia Volkensii* in the farms and the plantation in function of the diameter at breast height (DBH). The R^2 is mentioned next to the formulas.

Farms		
$D \text{ (cm)} = 61 + 31 \text{ dbh (cm)}$ (4.1)	$R^2 = 0.74$	
$h_1(m) = 2.1 + 0.43\text{dbh(cm)}$ (4.2)	$R^2 = 0.74$	
$h_2(m) = -1.4 + 0.76h_1(m)$ (4.3)	$R^2 = 0.83$	
Plantation		
$D \text{ (cm)} = 70 + 26 \text{ dbh (cm)}$ (4.4)	$R^2 = 0.59$	
$h_1(m) = 4.8 + 0.36\text{dbh(cm)}$ (4.5)	$R^2 = 0.41$	
$h_2(m) = -2.1 + 0.78h_1(m)$ (4.6)	$R^2 = 0.81$	

Radiometric parameters

Table 4.2 shows the estimated radiometric parameters used for the template generation algorithm. As expected, L_{max} shows the highest reflectance in the NIR band and the lowest in the blue band (Lillesand et al., 2015). The green band has a lower L_{max} than the red band. This could be explained because the tree mask (section 3.3.2) still contains background (BR) pixels, which have high reflectance in the red spectral band. The soil samples (BR) also contained some vegetation since crops and grasses are present between the trees, which can be noticed in the high NIR and green spectral reflectance values.

4. Results and discussion

Table 4.2: List of parameters used for the template generation algorithm for each feature. L_{max} : maximum radiance, amb: ambient, diffuse light and BR: background radiance.

Feature	L_{max}	amb	BR
Red	1315.667	727.200	1739.000
Green	1078.792	574.000	1189.182
Blue	829.450	417.750	894.727
NIR	2978.106	1645.050	2507.455
Brightness	1550.500	841.000	1582.600

4.1.2 Farms

The results of the optimisation process and final results in the farms can be found in table 4.3. The empirical and synthetic templates of the features resulting in the highest F-score* are visualised in figure 4.1. The complete template library for each feature can be found in appendix C.1 with further explanation about the used numbering. Figure 4.2 visualises the results for the features and template combinations resulting in the highest performance using the empirical and synthetic templates. The F-score of the brightness feature in the random combination is the highest during the training process, while the F-score* of the green feature in the random combination is highest for the final validation. Moreover, the difference is less than 0.006. For this reason, both were visualised and compared.

Table 4.3: Results of the template matching optimisation process for each feature with the optimal empirical (standard and random) and synthetic template combination and threshold. Both the overall performance during the training stage (F-score) and the final validation stage (F-score*) are displayed.

	Red	Green	Blue	NIR	Brightness
Standard					
Templates	34	28, 29	25, 26	19, 20	43
Threshold	0.668	0.501	0.603	0.855	0.450
recall	0.503	0.518	0.532	0.364	0.462
precision	0.479	0.470	0.554	0.250	0.296
F-score	0.491	0.493	0.543	0.297	0.361
F-score*	0.357	0.359	0.393	0.176	0.242
Random					
Templates	16, 40	13, 19	26, 37	17, 37	4, 7, 12
Threshold	0.761	0.668	0.591	0.834	0.702
recall	0.541	0.566	0.545	0.484	0.547
precision	0.562	0.579	0.484	0.482	0.615
F-score	0.552	0.573	0.513	0.483	0.579
F-score*	0.359	0.413	0.377	0.379	0.411
Synthetic					
Templates	3	1, 4	3, 4	1, 3, 5	3, 4, 5
Threshold	0.287	0.337	0.348	0.454	0.315
recall	0.378	0.125	0.152	0.222	0.081
precision	0.125	0.050	0.071	0.088	0.024
F-score	0.187	0.071	0.097	0.126	0.037
F-score*	0.129	0.061	0.075	0.062	0.044

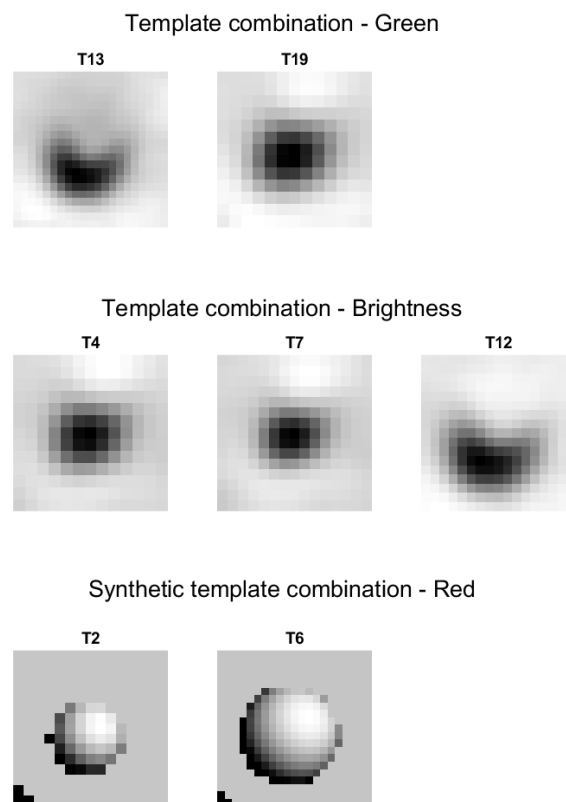


Figure 4.1: Template library of the farms of empirical (upper and middle) synthetic (lower) template combinations resulting in the highest performance for each feature.

4. Results and discussion

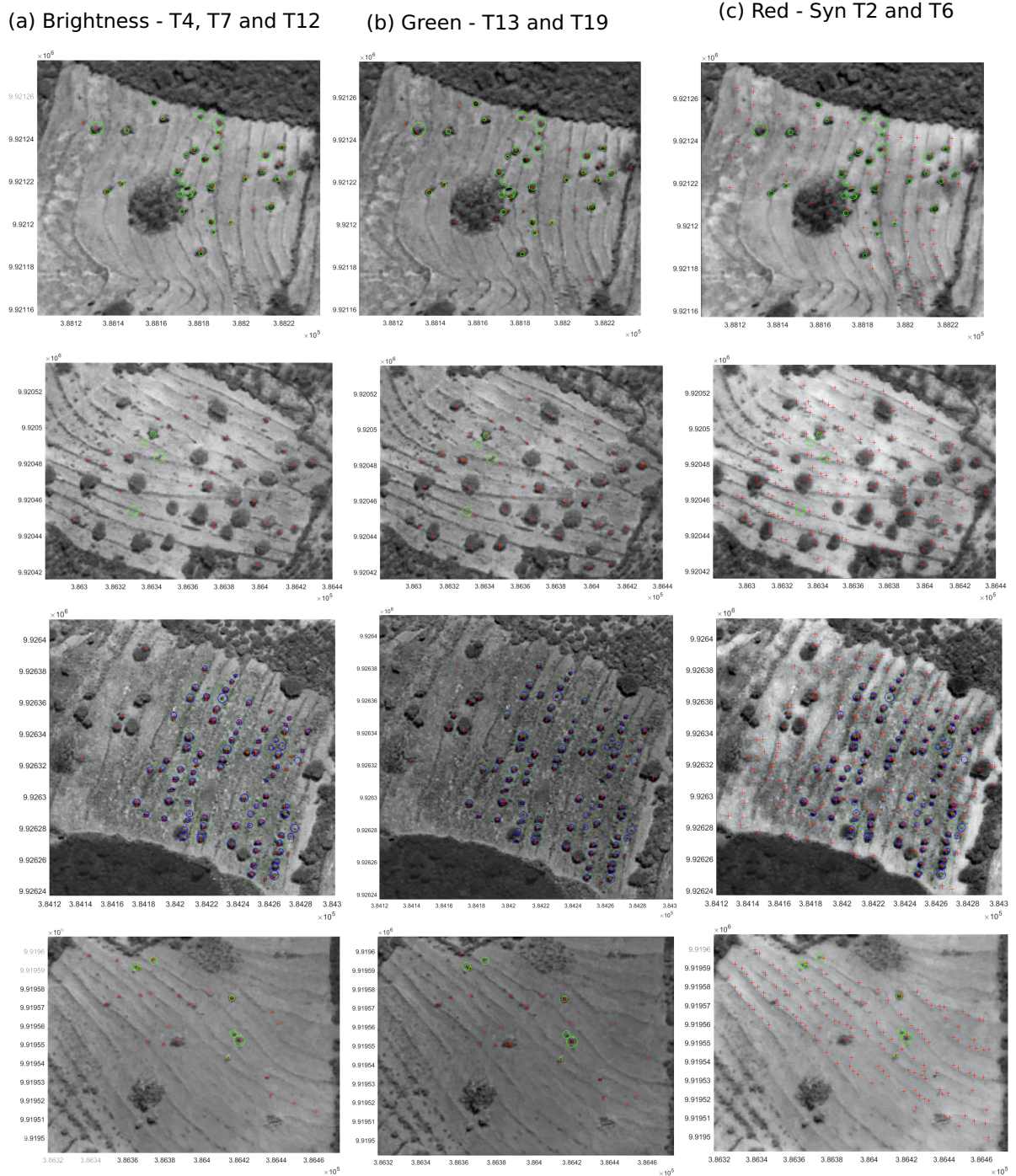


Figure 4.2: Visualisation of the detected trees (red) in farms 1, 5, 14 and 17 with the template matching algorithm for the feature brightness (column a), green (column b) and red (column c) with the template combination resulting in the highest performance as indicated in table 4.3. The ground truth validation (green) and training (blue) data are marked on the image. The size of the circles represents the crown diameter as measured during the field campaign.

4. Results and discussion

There are three remarks that can be made when analysing table 4.3. Firstly, the TM algorithm using the synthetic templates has a low recall and even lower precision than the empirical templates. This can also be observed in figure 4.2c, where the synthetic T3 template overdetects many trees by detecting false positives on the ridges of the farmlands. The optimal threshold values for the synthetic templates are much lower than those of the empirical templates, making them more susceptible to over detection (Hung et al., 2012) and could explain the lower precision. The reason for this low threshold is because algorithm 2 tries to find a balance between the recall and precision (section 2.3.2 and 3.4.1). If the simulated templates do not resemble the ground truth data well, it is hard for the TM algorithm to find a match between the image and a template, resulting in a low recall and thus, the threshold is lowered. It might be that the empirical templates correlate better with the *Melia volkensii* trees than the synthetic templates do. The shape function is approximated by formula 2.2, which assumes that the tree crowns are rotationally symmetric about the vertical axis (i.e., equal r_1 and r_2). Compared to figure 4.3, this seems to fit the trees not growing well under competition (i.e. the farms). Improving algorithm 3, by including more parameters, e.g. surface properties and leaf area distribution (Pollock, 1996) in the radiometric and geometric model (section 2.3.1), could lead to higher performance.



Figure 4.3: Picture of *Melia Volkensii* trees taken during the field visit in January 2022 in the plantation.

Secondly, it is also noticeable that including more templates does not always lead to higher performance: the green feature with T13 and T19 almost scored just as well as the brightness feature with T4, T7 and T12.

Lastly, for each feature, except for synthetic - brightness, the TM algorithm has a higher performance during the training stage than during the final validation. Because of this, the recall rate during the training stage (recall) and the recall rate during the final validation (recall*) were compared per diameter class in table 4.4. The percentage of training data and

4. Results and discussion

validation data per class are given in the last two columns. It was impossible to generate an overall performance per DBH class since false positives do not belong to any class. For the empirical templates, the less data available for a given class during the training stage, the higher the recall rate. This is not the case for the final validation, where the largest trees (the smallest data set) went undetected. For the synthetic templates, there is no pattern detected, except that the fifth class has the highest recall rate in both stages. It was expected that the highest recall rate would be observed in the classes where the corresponding crown diameter is more or less equal to the window size of the selected templates. The T4 empirical template has a window size of 13 x 13 pixels (6.5 m x 6.5 m). However, as is seen in figure 4.1, these templates contain a border of background pixels of approximately 2 pixels wide, making this template correspond with crown diameters of more or less 4.5 m. When comparing this crown diameter to table 3.3 this corresponds to diameter class 3. Following the same reasoning for the empirical T7 and T12, these templates correspond to DBH classes 4 and 6. No clear relation was found between the recall rate and the window size. The T3 synthetic template was simulated based on a tree with DBH of 15 cm, so this template belongs to class three's upper boundary and class four's lower boundary. However, these do not have the highest recall(*) rate.

Table 4.4: Recall rate per DBH class during the training stage and the final validation (*) for the empirical templates and the feature brightness and the synthetic template combination and the feature red. The right column shows the number of trees per class in the training data set and the validation data set.

DBH (cm)	recall		recall*		Percentage of trees (%)	
	Emperical	Synthethic	Emperical	Synthetic	training	validation
0.00 - 5.00	0.7011	0.322	0.333	0.410	16.76	16.74
5.00 - 10.00	0.6585	0.333	0.378	0.357	47.40	42.06
10.00 - 15.00	0.6752	0.319	0.103	0.400	30.25	24.89
15.00 - 20.00	0.7273	0.864	0.346	0.423	4.24	11.16
20.00 - 25.00	1	1	0.455	0.818	0.58	4.72
25.00 - 30.00	0.750	0.250	0	0	0.77	0.49

It is also important to note that the precision and, thereby, the performance are largely influenced by the accuracy of the vegetation mask generated by Forceville (2022). Forceville (2022) has concluded in the study that the tree mask has some shortcomings: some small trees and some trees on ridges were not included, while on the hand, other land cover classes and ridges containing crops (and no trees) were included. So it is impossible to conclude whether the precision is being over- or underestimated.

4.1.3 Plantation

The results of the optimisation process and final results can be found in table 4.5. The empirical and synthetic templates of the features resulting in the best PLA* are visualised in figure 4.4. The complete template library for each feature can be found in appendix C.2 with more explanation about the used numbering. Figure 4.5 visualises the results for each feature and template combination resulting in the best PLA* using the empirical and synthetic templates.

Table 4.5: Results of the template matching optimisation process for each feature with the optimal empirical (standard and random) and synthetic template combination and threshold. Both plot level accuracy (PLA) during the training stage and the final validation stage (PLA*) are displayed. A PLA greater than 1 indicates overdetetection and a PLA smaller than 1 indicates underdetetection.

	Red	Green	Blue	NIR	Brightness
Standard					
Templates	10, 11, 12	34	13, 14, 15	10, 11, 12	1, 2, 3
Threshold	0.058	0.005	0.047	0.200	0.505
PLA	2.576	2.358	2.640	2.511	2.731
PLA*	2.740	2.546	2.951	2.881	0.787
Random					
Templates	31, 37, 29	10, 31, 43	8, 15	43, 29, 4	27,15,43
Threshold	0.201	0.296	0.070	0.032	0.090
PLA	2.850	2.700	2.596	2.608	2.590
PLA*	2.909	2.694	2.810	2.943	2.842
Synthetic					
Templates	1, 4, 6	2, 3, 6	2, 4, 5	1, 2, 3	2, 6
Threshold	0.310	0.284	0.294	0.308	0.275
PLA	3.137	3.048	3.075	2.693	2.843
PLA*	3.339	3.083	3.056	3.344	2.967

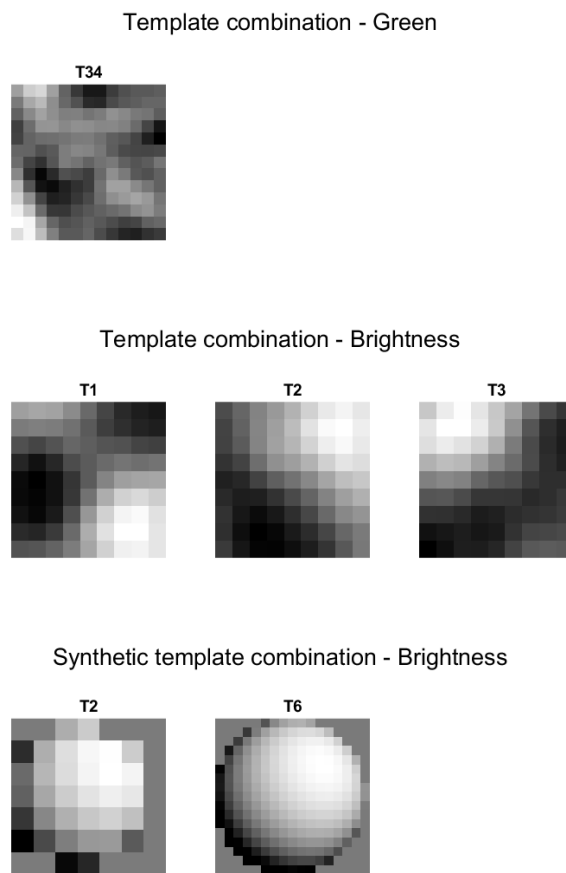


Figure 4.4: Template library of the plantation of the standard and random combinations resulting in the highest performance for each feature.

4. Results and discussion

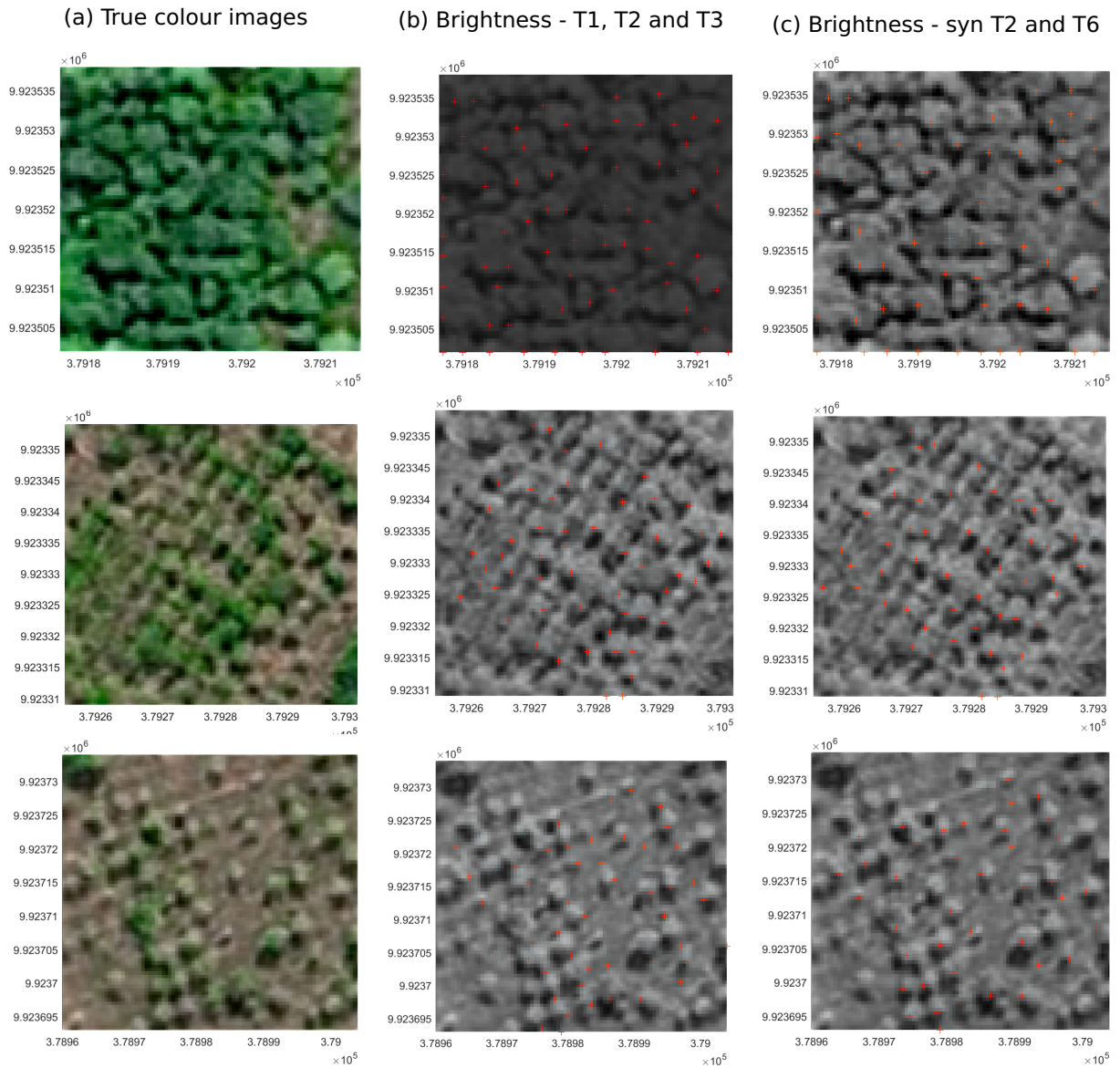


Figure 4.5: Visualisation of the detected trees (red) in the plantation with the template matching algorithm for the feature brightness for both the empirical (b) and synthetic (c) template combination resulting in the highest performance as indicated in table 4.5. The first row shows a dense plot, the second row an intermediate plot and the third a sparse plot.

4. Results and discussion

Two remarks can be made when analysing table 4.5. Firstly, in most cases, PLA is very high, meaning that the TM algorithm overdetects trees. Only the brightness feature with empirical T1, T2 and T3 underestimates the number of trees. However, when visually assessing the results, it looks like the detected trees have quite a good spatial accuracy and there is not as much over detection as the PLA predicts. There also seems to be no clear difference in performance between the plots with different tree densities. A reason for the incorrect PLA could be that the manual delineation of the plots was inaccurately done, causing an incorrect PLA.



Figure 4.6: Picture of *Melia Volkensii* trees taken during the field visit in January 2022 the plantation.

Secondly, when comparing the templates in figure 4.4, the tree in the synthetic and empirical templates do not resemble each other. The empirical templates have a better PLA (i.e. closer to 1). However, this is not apparent in figure 4.5, making it hard to draw a conclusion whether the empirical templates are an incorrect approximation of the trees or if the geometric parameters of the synthetic templates are incorrect. Both options are plausible since, on the one hand, the ground truth data sample has low spatial accuracy and on the other hand, many assumptions about the radiometric and geometric properties were made for the template generation. As previously stated, the shape function was approximated by equation 2.2, which assumes that the tree crowns are rotationally symmetric about the vertical axis (i.e., equal r_1 and r_2). Compared (figure 4.6), this shape function does not fit as well in closed stands in the plantation. Nonetheless, this function was applied to open and closed stands on the plantation. Moreover, to simulate the interaction of the tree crown with the scene irradiance, the model (equation 2.4) introduced by Gomes and Maillard (2014) was chosen, as described in section 2.3.1. It is important to note that this model assumes that the crown is an opaque surface. As stated in section 2.3.1, this is a rea-

sonable assumption for broad-leaved trees with leaves near the crown envelope (Pollock, 1996). However, this is a difficult assumption for mature trees growing under competition (i.e. the plantation). Still, this model was used in both layouts since no further knowledge about the leaf area distribution is available. Just as in the farms with the synthetic templates, the threshold is very low, making the template matching algorithm susceptible to over-detection (Hung et al., 2012) and suggesting that the templates do not correlate to the ground truth data very well.

4.2 Local maximum filtering

The results of the local maximum filtering approach are shown in table 4.6. The performance of the LMF was validated using the training data set and the validation data set so it could be compared to both stages of the TM algorithm. The features resulting in the highest F-score* and PLA* are visualised in figure 4.7b. The smoothed images are shown in figure 4.7b. The precision in the farms is very low, as can be seen in figure 4.7, where the whole background is detected as trees due to noise in the image. This result is expected as the method proposed by Khan and Gupta (2018) was developed for dense forests. After a visual assessment, the results on the plantation show better precision than the results on the farms. However, still, some trees are being detected multiple times. The trees in the plantation show a bright tree crown top, which is not present in the farms where the background has many bright pixels. A possible way to improve the LMF is by using a varying window size using the semivariance as proposed by Wulder et al. (2000) and described in section 2.4. However, it is not expected that this will improve the detection of trees in the farms since there are many non-forested areas that will still be detected using this method. Brandtberg et al. (2003) has mentioned that using an LMF where no trees are present will produce many false positives, which is always the case in farms. Implementing a method which performs edge detection prior to applying the LMF as proposed by Wang et al. (2004) and further elaborated in section 2.4 could work well in sparsely forested areas.

Table 4.6: Result for the farms and plantation for the features red, green, blue, NIR and brightness.

	Red	Green	Blue	NIR	Brightness
Farms					
recall	0.301	0.357	0.380	0.455	0.345
precision	0.010	0.012	0.013	0.016	0.012
F-score	0.020	0.023	0.025	0.030	0.023
F-score*	0.021	0.023	0.023	0.014	0.022
Plantation					
PLA	4.0069	4.003	4.000	4.016	3.959
PLA*	3.4544	3.457	3.440	3.494	3.425

4. Results and discussion

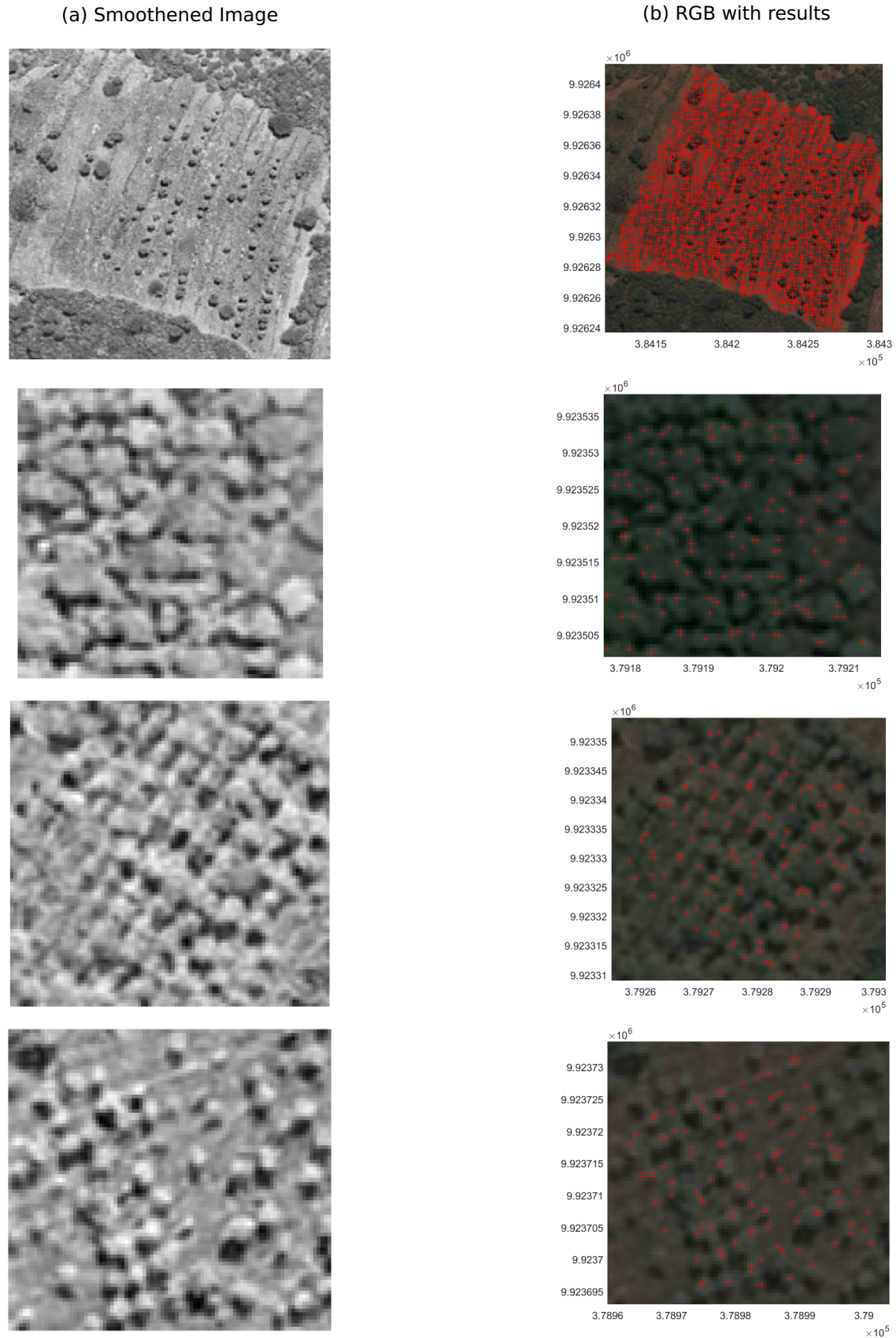


Figure 4.7: Results from the LMF algorithm within the first column the smoothened image and in the second column the results. The first row shows farm 14, the second row is a dense plot in the plantation, the third row is an intermediate dense plot in the plantation and the fourth row is a sparse plot in the plantation.

4.3 Method comparison

In the farms, the TM algorithm outperforms the LMF since, in the latter, the bright background pixels cause many false positives. The template matching algorithm also outperforms the LMF in the plantation since many crowns are detected multiple times. However, due to the inaccurate PLA, it is hard to draw a definite conclusion about all plots in the plantation since the difference is only subtle. This is less of a problem in the farms because the difference in performance is visually observed. When comparing the plantation versus the farms, the LMF performs better in the plantation. Again, it is hard to compare the performance of the TM algorithm in the farms and the plantation, but it seems that TM works well in both layouts.

Comparing the performance of TM with the synthetic and the empirical templates, it is apparent in the farms that the empirical templates perform better. The same statement can also be made with caution in the plantation. This means that there is still room for improvement of the template generation algorithm by optimising the geometric parameters in table 4.1.1, the radiometric parameters in table 3.4.1, the radiometric model (formula 2.4) and the geometric model (formula 2.2). Even though the synthetic tree model underperforms compared to the empirical template, it is still useful to optimise the template generation algorithm. Using a synthetic tree model has the benefit that no new ground truth data samples need to be made every time a new tree count is performed. Only the satellite image parameters (pixel size, sun elevation and sun azimuth) need to be adjusted, making it less labour-intensive.

A total tree count in both layouts was performed for a final comparison using both methods. The results can be seen in table 4.7. In all 17 farms, this result is 741 trees with the TM method and 25 775 with the LMF. The latter is an overestimation and the TM is the most reliable. During the field visit, 759 trees were counted (after excluding trees with a crown area smaller than 0.25 m²) in these 17 farms. Even though this is not a complete inventory, it gives an indication that the TM results are in the correct size order. The total count for the plantation is only performed on the non-cloud-covered areas, which comes down to 265 ha. This results in 450 trees/ha using the TM and 1120 trees/ha using the LMF. Assuming all trees in the plantation were planted with a planting distance of 4 m and no trees have died, this results in 625 trees/ha. This calculation again gives an indication that the TM results are more reliable.

Table 4.7: Total tree count performed in all 17 farms and the non-cloud-covered area in the plantation (265 ha) using the template matching method (TM) and the local maximum filtering method (LMF).

	Farms	Plantation
TM	741	119 274
LMF	25775	296 998

4.4 Recommendations

The proposed methods cannot accurately monitor the afforestation efforts of BGF. Since working with sub-par validation and training data and satellite imagery and only simplified ITCD algorithms are used. In further developing a sustainable monitoring tool, some adjustments can be made, leading to higher performance. Some recommendations are made on how to improve the data collection (section 4.4.1) during a field survey, the algorithm optimisation (section 4.4.2) and the used imagery (4.4.3).

4.4.1 Data collection

The inaccurate location of the training data was one of the most demanding challenges during this study, complicating the template matching training stage (section 3.4) and the validation of the detected trees on the plantation. The data points in the farms were manually corrected by Forceville (2022); however, this was a very time-consuming process. The handheld Garmin 60 csx global positioning system (GPS) had a positional accuracy of three meters at best, which is proven to be insufficient. Nilsson et al. (2017) have used a high-precision GPS receiver, the Spectra Precision MobileMapper 120, giving a positional accuracy of 1 m. Moreover, between 2024 and 2030, the African satellite augmentation system (ASAS) will be available in Kenya (Ilčev, 2018). GPS devices compatible with this system will have improved positional accuracy (Forceville, 2022; Ilčev, 2018). Investing in a high-precision GPS will be not only beneficial for the development of the monitoring tool but also valuable for the long run for the monitoring itself since, in the farms, it will be necessary to register the coordinates of trees when being planted. This is because farmers grow their *Melia Volkensii* trees among the trees planted by BGF, making it impossible to distinguish which one belongs to who. This will help to monitor whether a tree planted by BGF in the farms is still alive after one or more growing seasons and how much it has grown. This problem does not occur on the plantation because all trees are being planted by BGF.

In addition to the more accurate tree locations, more full inventories of the sample plots in the plantation need to be made as well to improve the validation stage of the methods. Since in the field survey often trees were not measured because they were different from the trees that were planted by BGF e.g. the baobab trees. However, it is necessary to measure all trees in the plots so the true negatives and false positives can be calculated without the need of a vegetation mask. Furthermore, the coordinates of the outlines of the sample plots in the plantation also need to be stored so they can be located more precisely on the image. The outline of the plots was manually determined, which introduced additional errors to the validation of the obtained tree points.

4.4.2 Algorithm optimisation

The additional data which needs to be collected, as described above, will aid the training stage of the TM algorithm. When more positional accurate training data is available, the determination of the optimal template combination, the best feature and the optimal threshold can be more precise. When this data is available better conclusion can be made

about the final results since now, in both the plantation and the farms, these need to be cautiously interpreted. By obtaining more data about the trees, possible better relations between the diameter at breast height and the height, crown height, and crown diameter could be found, resulting in better templates. Moreover, different geometrical models should be used on farms and plantations. As described in section 4.1.3, the used geometrical model, a round sphere, does not fit the trees in dense plots in the plantation well. The synthetic trees can also be improved by refining the radiometric model used for the template generation since now many assumptions have been made (Pollock, 1996). TM using the synthetic templates could also be improved by accounting for more background variations: bare soil and soil planted with crops. In further research, using a larger template library and applying more than three templates can be further explored. Also testing out other features such as the normalized difference vegetation index (NDVI), the first principal component of a principal component analysis (PCA) (Wang et al., 2004) and the ratio between a red edge band and a red band (Bunting and Lucas, 2006) can be considered. To find the optimal threshold resulting in the best performance (i.e. balancing over-detection and under-detection), the threshold is increased from 0 to 1 with 0.1 per iteration. Taking a smaller stepsize and, thus, more iterations will lead to a more accurate threshold determination. However, increasing the number of templates, features and the number of iterations will increase the computation time, making the algorithm's training stage more time-consuming. Finding the optimal threshold (number of iterations = 10) and the optimal template combination (maximum three templates and number of iterations = 100) took about four hours per feature. Another possibility is to explore the method proposed by Maillard and Gomes (2016), which combines template matching with a marked point process and see if the performance increases.

It is possible to improve the local maximum filtering process by using a varying window size using the semivariance as proposed by Wulder et al. (2000) and described in section 2.4. The pre-processing method was proposed by Khan and Gupta (2018) and developed for dense forests. Other methods, such as the method proposed by Wang et al. (2004), which performs an edge detection before applying a local maximum filter, could be explored in further research. However, Brandtberg et al. (2003) has mentioned that using an LMF where no trees are present will produce many false positives, which is always the case in farms. For this reason, it is recommended to only apply the LMF to densely forested areas in further research.

4.4.3 Satellite and aerial imagery

To detect the smallest trees, the scale-space method can be further explored: detecting trees of different sizes in the same grayscale image at different scale levels (Larsen et al., 2011) and using satellite and aerial imagery with a higher spatial resolution so a higher crown-to-pixel size ratio can be obtained for the smallest trees (Pouliot et al., 2002) (section 2.1). It is expected that template matching will perform better on cm-level imagery. Additional to satellite imagery, other data sources can be further explored. Unmanned Aerial Vehicles (UAVs) are a technology that can provide high-resolution information, for example, 20 cm/pixel, over small areas cost-effectively (Hung et al., 2012). Dainelli et al. (2021) has mentioned that when UAVs are appropriately used, they can constitute a valu-

able tool in monitoring and mapping forests. Many studies have used aerial images in combination with TM and LMF, e.g. Larsen and Rudemo (1998), Pouliot et al. (2002), Erikson (2004). Kempf et al. (2021) has used oblique view aerial images to generate a digital surface model for tree detection. Zhen et al. (2016) has mentioned that combining passive and active remotely sensed data may increase performance of ITCD. Many studies have developed methods which incorporate Light Detection And Ranging (LiDAR) data with UAVs which can be used- to collect high-resolution imagery and LiDar data. However, it is also important to note that using UAVs has some drawbacks related to technical issues (Dainelli et al., 2021). Due to battery duration, the data cannot be effectively acquired for the whole extent of a large area, which is the case for BGF. There are also limitations on airspace use enforced by policy and regulations (Dainelli et al., 2021).

5. CONCLUSION

Is it possible to accurately count Melia Volkensii trees using very high resolution (VHR) satellite imagery? In this thesis, two methods were proposed for detecting *Melia Volkensii* trees in an agroforestry layout (i.e. farms) and in a plantation in a study area in Kiambere. Unfortunately, even though the methods show promising results, in this study, the trees cannot be accurately counted with the PLéiades satellite image.

Is there a different method needed for the different afforestation layouts? The first method, template matching (TM), uses geometrical-optical tree models based on training data (i.e. empirical templates) or based on illumination and radiance parameters and tries to find where this model matches the image using the normalised cross correlation similarity measure. The TM algorithm takes the trees' radiometric and geometric properties into consideration (Wang et al., 2004) and can exploit contextual information such as the shadow cast by trees (Maillard and Gomes, 2016) and soil reflection. Before matching the templates to the image, the input parameters, namely the greyscale image (i.e., which feature to use), the threshold and the template combination algorithm, were optimised. The TM algorithm shows promising results in both layouts, with a performance of 0.411 in the farms with the brightness feature and empirical templates T4, T7 and T12 and a plot level accuracy (PLA) of 0.787 in the plantation with the brightness feature and empirical templates T1, T2 and T3. Unfortunately, due to the lack of spatially accurate ground truth data, no performance in the plantation could be calculated. Moreover, in the plantation, the PLA must be interpreted cautiously since the plots were manually delineated and could contain errors. The synthetic templates underperform compared to the empirical templates in both layouts. However, optimising the template generation algorithm is still useful since using a synthetic tree model has the benefit that no new ground truth data samples need to be made every time a new tree count is performed. Only the satellite image parameters (pixel size, sun elevation and sun azimuth) need to be adjusted, making it less labour-intensive. The second method, local maximum filtering, is a technique that assumes that points with the highest brightness (i.e. digital number) within a search window are possible tree locations (Pouliot et al., 2002; Quackenbush et al., 2000). Mathematical morphology procedures (opening, closing, opening/closing by reconstruction) were used to filter the image before applying a non-overlapping local maximum filter. The LMF performs sub-par compared to TM, with only a performance of 0.023 in the farms using the green feature and a plot level accuracy of 3.425 with the brightness feature. So to conclude: there is no different method needed in the different afforestation layouts since TM performs well in both.

Which method has the highest performance? In this particular case study, the template matching approach performs superior to the LMF. However, further development and research are needed to optimise and improve the algorithm to accurately detect *Melia Volkensii* in both layouts.

5. Conclusion

To accurately monitor the afforestation efforts of Better Globe Forestry, more spatial accurate tree positions are needed, the TM algorithm needs further optimisation and higher resolution imagery is needed.

BIBLIOGRAPHY

- Airbus (2011). *Pléiades Imagery – User Guide*. Airbus.
- Alimuddin, I., Sumantyo, J. T. S., Kuze, H., et al. (2012). Assessment of pan-sharpening methods applied to image fusion of remotely sensed multi-band data. *International Journal of Applied Earth Observation and Geoinformation*, 18:165–175.
- Benediktsson, J. A., Pesaresi, M., and Amason, K. (2003). Classification and feature extraction for remote sensing images from urban areas based on morphological transformations. *IEEE transactions on geoscience and remote sensing*, 41(9):1940–1949.
- Benz, U. C., Hofmann, P., Willhauck, G., Lingenfelder, I., and Heynen, M. (2004). Multi-resolution, object-oriented fuzzy analysis of remote sensing data for gis-ready information. *ISPRS Journal of photogrammetry and remote sensing*, 58(3-4):239–258.
- BGF (n.d.). Better globe forestry - about us.
- Biswas, H., Zhang, K., Ross, M. S., and Gann, D. (2020). Delineation of tree patches in a mangrove-marsh transition zone by watershed segmentation of aerial photographs. *Remote Sensing*, 12(13):2086.
- Brandtberg, T. and Walter, F. (1998). Automated delineation of individual tree crowns in high spatial resolution aerial images by multiple-scale analysis. *Machine Vision and Applications*, 11(2):64–73.
- Brandtberg, T. and Warner, T. (2006). High-spatial-resolution remote sensing. In *Computer applications in sustainable forest management*, pages 19–41. Springer.
- Brandtberg, T., Warner, T. A., Landenberger, R. E., and McGraw, J. B. (2003). Detection and analysis of individual leaf-off tree crowns in small footprint, high sampling density lidar data from the eastern deciduous forest in north america. *Remote sensing of Environment*, 85(3):290–303.
- Bunting, P. and Lucas, R. (2006). The delineation of tree crowns in australian mixed species forests using hyperspectral compact airborne spectrographic imager (casi) data. *Remote Sensing of Environment*, 101(2):230–248.
- Cheng, G. and Han, J. (2016). A survey on object detection in optical remote sensing images. *ISPRS Journal of Photogrammetry and Remote Sensing*, 117:11–28.
- Cook, K. H. and Vizzy, E. K. (2013). Projected changes in east african rainy seasons. *Journal of Climate*, 26(16):5931–5948.
- Culvenor, D. S. (2000). Development of a tree delineation algorithm for application to high spatial resolution digital imagery of australian native forest.

- Culvenor, D. S. (2002). Tida: an algorithm for the delineation of tree crowns in high spatial resolution remotely sensed imagery. *Computers & Geosciences*, 28(1):33–44.
- Curran, P. J. and Atkinson, P. M. (1998). Geostatistics and remote sensing. *Progress in Physical Geography*, 22(1):61–78.
- Dainelli, R., Toscano, P., Di Gennaro, S. F., and Matese, A. (2021). Recent advances in unmanned aerial vehicle forest remote sensing—a systematic review. part i: A general framework. *Forests*, 12(3):327.
- Daley, N. M., Burnett, C. N., Wulder, C., Niemann, K. O., and Goodenough, D. G. (1998). Comparison of fixed-size and variable-sized windows for the estimation of tree crown position. In *IGARSS'98. Sensing and Managing the Environment. 1998 IEEE International Geoscience and Remote Sensing. Symposium Proceedings*. (Cat. No. 98CH36174), volume 3, pages 1323–1326. IEEE.
- Descombes, X. and Pechersky, E. (2006). *Tree Crown Extraction using a Three State Markov Random Field*. PhD thesis, INRIA.
- Ehlers, M., Klonus, S., Johan Åstrand, P., and Rosso, P. (2010). Multi-sensor image fusion for pansharpening in remote sensing. *International Journal of Image and Data Fusion*, 1(1):25–45.
- Erikson, M. (2003). Segmentation of individual tree crowns in colour aerial photographs using region growing supported by fuzzy rules. *Canadian Journal of Forest Research*, 33(8):1557–1563.
- Erikson, M. (2004). *Segmentation and classification of individual tree crowns*, volume 320.
- Erikson, M. (2006). Two preprocessing techniques based on grey level and geometric thickness to improve segmentation results. *Pattern Recognition Letters*, 27(3):160–166.
- Erikson, M. and Olofsson, K. (2005). Comparison of three individual tree crown detection methods. *Machine Vision and Applications*, 16(4):258–265.
- Forceville, E. (2022). Monitoring afforestation with very high resolution satellite imagery in Kenya: a method to differentiate *Melia Volkensii*. Master's thesis, Ghent University.
- Fouda, Y. (2014). Template matching from 2-d into 1-d. *Journal of Signal and Information Processing*, 05:23–31.
- Gomes, M. F. and Maillard, P. (2013). Identification of urban tree crown in a tropical environment using worldview-2 data: Problems and perspectives. In *Earth Resources and Environmental Remote Sensing/GIS Applications IV*, volume 8893, pages 60–72. SPIE.
- Gomes, M. F. and Maillard, P. (2014). Integration of marked point processes and template matching for the identification of individual tree crowns in an urban and a wooded savanna environment in brazil. In *Earth Resources and Environmental Remote Sensing/GIS Applications V*, volume 9245, pages 220–231. SPIE.
- Gomes, M. F., Maillard, P., and Deng, H. (2018). Individual tree crown detection in sub-meter satellite imagery using marked point processes and a geometrical-optical model. *Remote Sensing of Environment*, 211:184–195.

- Gong, P., Sheng, Y., and Biging, G. (2002). 3d model-based tree measurement from high-resolution aerial imagery. *Photogrammetric engineering and remote sensing*.
- Gonzalez, R. C. and Woods, R. E. (2008). *Digital image processing*. Prentice Hall, Upper Saddle River, N.J.
- Gougeon, F. (1997). A locally adaptive technique for forest regeneration assessments from high resolution aerial images.
- Gougeon, F. A. (1995). A crown-following approach to the automatic delineation of individual tree crowns in high spatial resolution aerial images. *Canadian journal of remote sensing*, 21(3):274–284.
- Heipke, C. et al. (2007). Automatic extraction and delineation of single trees from remote sensing data. *Machine Vision and Applications*, 18(5):317–330.
- Hirschmugl, M., Ofner, M., Raggam, J., and Schardt, M. (2007). Single tree detection in very high resolution remote sensing data. *Remote Sensing of Environment*, 110(4):533–544.
- Hisham, M., Yaakob, S. N., Raof, R., Nazren, A. A., and Wafi, N. (2015). Template matching using sum of squared difference and normalized cross correlation. In *2015 IEEE Student Conference on Research and Development (SCoReD)*, pages 100–104.
- Hung, C., Bryson, M., and Sukkarieh, S. (2012). Multi-class predictive template for tree crown detection. *ISPRS journal of photogrammetry and remote sensing*, 68:170–183.
- Ilčev, S. D. (2018). Global satellite augmentation systems (gsas). In *Global Mobile Satellite Communications Applications*, pages 467–559. Springer.
- Jain, A. K., Zhong, Y., and Dubuisson-Jolly, M.-P. (1998). Deformable template models: A review. *Signal processing*, 71(2):109–129.
- Kalapala, M. (2014). Estimation of tree count from satellite imagery through mathematical morphology. *International Journal of Advanced Research in Computer Science and Software Engineering*, 4(1):490–495.
- Kaszta, Ž., Van De Kerchove, R., Ramoelo, A., Cho, M. A., Madonsela, S., Mathieu, R., and Wolff, E. (2016). Seasonal separation of african savanna components using worldview-2 imagery: a comparison of pixel-and object-based approaches and selected classification algorithms. *Remote Sensing*, 8(9):763.
- Ke, Y. and Quackenbush, L. J. (2011). A review of methods for automatic individual tree-crown detection and delineation from passive remote sensing. *International Journal of Remote Sensing*, 32(17):4725–4747.
- Kempf, C., Tian, J., Kurz, F., D’Angelo, P., Schneider, T., and Reinartz, P. (2021). Oblique view individual tree crown delineation. *International Journal of Applied Earth Observation and Geoinformation*, 99:102314.
- Khan, S. and Gupta, P. K. (2018). Comparative study of tree counting algorithms in dense and sparse vegetative regions. *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*.
- Klonus, S. and Ehlers, M. (2009). Performance of evaluation methods in image fusion. In *2009 12th International Conference on Information Fusion*, pages 1409–1416. IEEE.

- Korpela, I., Dahlin, B., Schäfer, H., Bruun, E., Haapaniemi, F., Honkasalo, J., Ilvesniemi, S., Kuutti, V., Linkosalmi, M., Mustonen, J., et al. (2007). Single-tree forest inventory using lidar and aerial images for 3d treetop positioning, species recognition, height and crown width estimation. In *Proceedings of ISPRS workshop on laser scanning*, pages 12–14.
- Korpela, I. et al. (2004). Individual tree measurements by means of digital aerial photogrammetry.
- Lamar, W. R., McGraw, J. B., and Warner, T. A. (2005). Multitemporal censusing of a population of eastern hemlock (*tsuga canadensis* L.) from remotely sensed imagery using an automated segmentation and reconciliation procedure. *Remote Sensing of Environment*, 94(1):133–143.
- Larsen, M. (1997). Crown modelling to find tree top positions in aerial photographs. In *Third International Airborne Remote Sensing Conference and Exhibition*, volume 7, page 10. Citeseer.
- Larsen, M., Eriksson, M., Descombes, X., Perrin, G., Brandtberg, T., and Gougeon, F. A. (2011). Comparison of six individual tree crown detection algorithms evaluated under varying forest conditions. *International Journal of Remote Sensing*, 32(20):5827–5852.
- Larsen, M. and Rudemo, M. (1998). Optimizing templates for finding trees in aerial photographs. *Pattern Recognition Letters*, 19(12):1153–1162.
- Lévesque, J. and King, D. J. (1999). Airborne digital camera image semivariance for evaluation of forest structural damage at an acid mine site. *Remote Sensing of Environment*, 68(2):112–124.
- Lillesand, T., Kiefer, R. W., and Chipman, J. (2015). *Remote sensing and image interpretation*. John Wiley & Sons.
- Lindeberg, T. (1996). Scale-space: A framework for handling image structures at multiple scales.
- Liu, J., Heiskanen, J., Aynekulu, E., and Pellikka, P. (2015). Seasonal variation of land cover classification accuracy of landsat 8 images in burkina faso. *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*.
- Maillard, P. and Gomes, M. F. (2016). Detection and counting of orchard trees from vhr images using a geometrical-optical model and marked template matching. *ISPRS Annals of Photogrammetry, Remote Sensing & Spatial Information Sciences*, 3(7).
- Manning, C., Raghavan, P., and Schütze, H. (2008). *An Introduction to Information Retrieval*, volume 1. Cambridge University Press.
- Mather, P. and Tso, B. (2016). *Classification methods for remotely sensed data*. CRC press.
- Minnaert, M. (1941). The reciprocity principle in lunar photometry. *The Astrophysical Journal*, 93:403–410.
- Mishra, N. B. and Crews, K. A. (2014). Mapping vegetation morphology types in a dry savanna ecosystem: Integrating hierarchical object-based image analysis with random forest. *International Journal of Remote Sensing*, 35(3):1175–1198.

- Nilsson, M., Nordkvist, K., Jonzén, J., Lindgren, N., Axensten, P., Wallerman, J., Egberth, M., Larsson, S., Nilsson, L., Eriksson, J., et al. (2017). A nationwide forest attribute map of sweden predicted using airborne laser scanning data and field data from the national forest inventory. *Remote Sensing of Environment*, 194:447–454.
- Omuto, C. (2008). Assessment of soil physical degradation in eastern kenya by use of a sequential soil testing protocol. *Agriculture, Ecosystems I& Environment*, 128:199–211.
- Ozendi, M., Topan, H., Oruc, M., and Cam, A. (2016). Pan-sharpening quality investigation of pléiades-1a images. *Geocarto International*, 31(8):881–890.
- Pollock, R. (1996). *The automatic recognition of individual trees in aerial images of forests based on a synthetic tree crown image model*. PhD thesis, University of British Columbia.
- Pouliot, D. and King, D. (2005). Approaches for optimal automated individual tree crown detection in regenerating coniferous forests. *Canadian Journal of Remote Sensing*, 31(3):255–267.
- Pouliot, D., King, D., Bell, F., and Pitt, D. (2002). Automated tree crown detection and delineation in high-resolution digital camera imagery of coniferous forest regeneration. *Remote sensing of environment*, 82(2-3):322–334.
- Pratt, W. K. (1991). *Digital image processing*. John Wiley Sins, Inc.
- Pu, R., Landry, S., and Yu, Q. (2018). Assessing the potential of multi-seasonal high resolution pléiades satellite imagery for mapping urban tree species. *International Journal of Applied Earth Observation and Geoinformation*, 71:144–158.
- Quackenbush, L. J., Hopkins, P. F., and Kinn, G. J. (2000). Using template correlation to identify individual trees in high resolution imagery. In *Proceedings of the 2000 ASPRS Annual Conference*, pages 22–26.
- Salamí, E., Gallardo, A., Skorobogatov, G., and Barrado, C. (2019). On-the-fly olive trees counting using a uas and cloud services. *Remote Sensing*, 11:316.
- Shen, X. and Bao, W. (2014). The remote sensing image matching algorithm based on the normalized cross-correlation and sift. *Journal of the Indian Society of Remote Sensing*, 42(2):417–422.
- Soille, P. et al. (1999). *Morphological image analysis: principles and applications*, volume 2. Springer.
- Terol-Villalobos, I. R. and Vargas-Vázquez, D. (2005). Openings and closings with reconstruction criteria: a study of a class of lower and upper levelings. *Journal of Electronic imaging*, 14(1):013006.
- Trimble (2022). *Trimble Documentation eCognition Developer 10.2.2 User Guide*. Trimble Germany GmbH: Munich.
- Vahidi, H., Klinkenberg, B., Johnson, B. A., Moskal, L. M., and Yan, W. (2018). Mapping the individual trees in urban orchards by incorporating volunteered geographic information and very high resolution optical remotely sensed data: A template matching-based approach. *Remote Sensing*, 10(7):1134.

- Van Coillie, F. M., Verbeke, L. P., and De Wulf, R. R. (2007). Feature selection by genetic algorithms in object-based classification of ikonos imagery for forest mapping in flanders, belgium. *Remote Sensing of Environment*, 110(4):476–487.
- Vandenabeele, J. (2006-2021). [monthly rainfall in kiambere][unpublished raw data].
- Venter, F. J., Scholes, R. J., and Eckhardt, H. C. (2003). The abiotic template and its associated vegetation pattern. *The Kruger experience: Ecology and management of savanna heterogeneity*, 83:129.
- Wang, L., Gong, P., and Biging, G. S. (2004). Individual tree-crown delineation and tree-top detection in high-spatial-resolution aerial imagery. *Photogrammetric Engineering & Remote Sensing*, 70(3):351–357.
- Woodcock, C. E., Strahler, A. H., and Jupp, D. L. (1988). The use of variograms in remote sensing: I. scene models and simulated images. *Remote Sensing of Environment*, 25(3):323–348.
- Wulder, M., Niemann, K. O., and Goodenough, D. G. (2000). Local maximum filtering for the extraction of tree locations and basal area from high spatial resolution imagery. *Remote Sensing of environment*, 73(1):103–114.
- Xiao, C., Qin, R., Huang, X., and Li, J. (2018). Individual tree detection from multi-view satellite images. In *IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium*, pages 3967–3970. IEEE.
- Zhen, Z., Quackenbush, L. J., and Zhang, L. (2016). Trends in automatic individual tree crown detection and delineation—evolution of lidar data. *Remote Sensing*, 8(4):333.

APPENDIX A

LOCATIONS OF THE DATA

COLLECTION

Table A.1: Location of the corners of the samples plot registered in WGS 84 (Worldwide Geodetic System 84) datum and the Universal Transverse Mercator (UTM) zone 37S projection.

No.	East (UTM)	South (UTM)	No.	East (UTM)	South (UTM)
1	378648.255	9925451.187	23	379315.064	9924069.307
2	379216.417	9925145.643	23	379402.699	9924079.318
3	379353.029	9925153.344	24	378780.631	9923663.785
4	378832.919	9924920.164	24	378808.354	9923739.868
4	378703.547	9925004.565	25	378982.699	9923712.607
5	379034.833	9924932.024	25	379224.965	9923860.000
5	379115.537	9924931.099	26	379320.455	9923706.447
6	379408.320	9924933.410	26	379470.466	9923782.838
6	379607.000	9925038.448	27	378861.797	9923560.440
7	379745.922	9924933.102	28	379192.468	9923520.550
8	378439.256	9924812.354	28	379011.038	9923494.060
10	379097.518	9924818.514	29	379512.512	9923444.775
10	378929.641	9924767.843	29	379313.986	9923498.834
11	379619.860	9924678.745	30	378545.065	9923266.117
11	379312.600	9924677.821	31	378925.868	9923148.835
12	379825.394	9924693.685	31	379278.409	9923332.498
13	378832.765	9924508.173	32	379429.883	9923188.031
13	378690.609	9924478.757	34	378795.340	9922944.302
14	379133.095	9924439.175	35	379009.883	9923043.180
14	379047.771	9924513.410	36	379631.489	9923008.604
15	379492.952	9924524.422	36	379580.818	9923005.369
15	379491.874	9924596.963	37	378339.685	9922728.912
16	379740.223	9924596.501	38	378576.484	9922785.281
17	378879.278	9924186.590	38	378531.203	9922809.154
18	379248.992	9924251.661	39	378924.712	9922713.356
18	379039.300	9924147.932	40	379453.293	9922811.233
19	379329.080	9924293.861	41	378476.759	9922418.879
19	379405.009	9924236.722	42	378728.189	9922552.256
20	379725.823	9924356.776	45	378438.640	9922260.012
21	378882.897	9923940.242	46	378568.013	9922269.253
22	379004.877	9924051.287	46	378600.125	9922196.250
22	379097.287	9924064.532	49	378596.737	9922076.272

A. Locations of the data collection

Table A.2: Location of the farms registered in WGS 84 (Worldwide Geodetic System 84) datum and the Universal Transverse Mercator (UTM) zone 37S projection. The village and name of the farmer are mentioned in the last two columns.

No.	East (UTM)	South (UTM)	Village	Name of farmer
1	381295.869	9925051.410	Katithini	Stephene Mulwa Mukiti
2	381118.480	9924378.304	Katithini	Peter Muimi Muthui
3	381653.000	9923907.733	Katithini	Peter Mutua Gitune
4	384629.458	9922789.833	Mwangu	Muteti Muthui
5	386408.941	9921059.196	Maskanioni	Bonface Mutua Musyoka
6	386404.878	9919557.809	Masukanioni	Peter Munyoki Muthui
7	386873.694	9920308.766	Masukanioni	Peter Kimanzi Makau
8	388171.355	9921223.529	Masukanioni	Peter mwendwa Francis mutemi
9	386361.811	9920473.338	Masukanioni	Samuel musili munyoki
10	390655.912	9924837.998	Koriro	Benson Kimanzi Mwendwa
11	388753.026	9925307.343	Kalatine	Mboli Ngukuni
12	388313.971	9925142.750	Kalatine	Martha Mwendwa
13	384217.805	9926323.223	Tondora	Kilonzi Kimwele Ndatya
14	383473.495	9927602.109	Tondora	Muthakye Muteti
15	382890.209	9926853.099	Katithini	Peter Mutinda Muthengi
16	381305.399	9925882.234	Katithini	Muthui Masyuko
17	381754.064	9924579.663	Katithini	Mutuku Kitune

APPENDIX B

DESCRIPTIVE STATISTICS

Table B.1: Descriptive statics of *Melia Volkensii* (Mukau), *Azadirachta indica* (Neem) and *Adansonia sp.* (Baobab). The trees of Mukau and Neem are divided into diameter classes of 5 cm based on the diameter at breast height (DBH). Since only 3 Baobabs were measured, these trees were divided into 2 classes. Per class the table contains the average value of the crown diameter, the height, the tree age and the tree vitality. The last column shows the percentage of trees per diameter compared to the total amount of that species.

DBH (cm)	Crown diameter (cm)	Height (cm)	Tree age (year)	Tree Vitality	Number of trees (%)
<i>Azadirachta indica</i>					
0.0 - 5.0	155.46	3.34	14.00	2.60	13.87
5.0 - 10.0	251.61	5.39	14.00	1.94	46.72
10.0 - 15.0	337.15	7.90	13.90	1.42	32.36
15.0 - 20.0	386.88	9.52	13.85	1.58	6.33
20.0 - 25.0	595.17	9.57	14.00	1.00	0.73
<i>Acacia sp.</i>					
0.0 - 5.0	397.75	2.50	10.00	2.00	2.50
5.0 - 10.0	462.26	6.84	10.00	1.96	6.57
10.0 - 15.0	612.06	10.75	10.0	1.09	8.27
15.0 - 20.0	664.66	11.44	10.00	1.00	3.89
20.0 - 25.0	776.50	9.60	10.00	1.00	0.24
<i>Adansonia sp.</i>					
0.0 - 50.0	280.00	4.00	30.00	1.00	66.67
350.0 - 400.0	2377.50	21.30	600.00	1.00	33.33

APPENDIX C

TEMPLATE LIBRARY

The numbering of the empirical templates have the following meaning for the farms as well as the plantation:

- T1 - T3: window size 9 with 3 subgroups.
- T4 - T6: window size 13 with 3 subgroups.
- T7 - T9: window size 15 with 3 subgroups.
- T10 - T12: window size 17 with 3 subgroups.
- T13 - T15: window size 21 with 3 subgroups.
- T16 - T18: window size 9 with 2 subgroups and one empty template.
- T19 - T21: window size 13 with 2 subgroups and one empty template.
- T22 - T24: window size 15 with 2 subgroups and one empty template.
- T25 - T27: window size 17 with 2 subgroups and one empty template.
- T28 - T30: window size 21 with 2 subgroups and one empty template.
- T31 - T33: window size 9 with no subgroups and two empty template.
- T34 - T36: window size 13 with no subgroups and two empty template.
- T37 - T39: window size 15 with no subgroups and two empty template.
- T40 - T42: window size 17 with no subgroups and two empty template.
- T43 - T45: window size 21 with no subgroups and two empty template.

A template was generated for each upper limit of the diameter classes in table 3.3 for the farms. The other tree parameters needed for the template generation algorithm were estimated using equations 4.1, 4.4 and 4.2, table 3.4.1 and a buffer value of 3. The template numbers in the farms have the following diameter at breast height:

- T1: DBH = 5 cm
- T2: DBH = 10 cm
- T3: DBH = 15 cm
- T4: DBH = 20 cm

C. Template library

- T5: DBH = 25 cm
- T6: DBH = 30 cm
- T7: empty template
- T8: empty template
- T9: empty template

A template was generated for each upper limit of the diameter classes in table 3.3 for the plantation. The other tree parameters needed for the template generation algorithm were estimated using equations 4.5, 4.3 and 4.6, table 3.4.1 and a buffer value of 0. The template numbers in the plantation have the following diameter at breast height:

- T1: DBH = 5 cm
- T2: DBH = 10 cm
- T3: DBH = 15 cm
- T4: DBH = 20 cm
- T5: DBH = 25 cm
- T6: DBH = 30 cm
- T7: DBH = 30 cm
- T8: DBH = 30 cm
- T9: DBH = 30 cm
- T10: DBH = 30 cm
- T11: empty template
- T12: empty template
- T13: empty template

The resulting templates can be found in section C.1 and C.2.

C.1 Farms

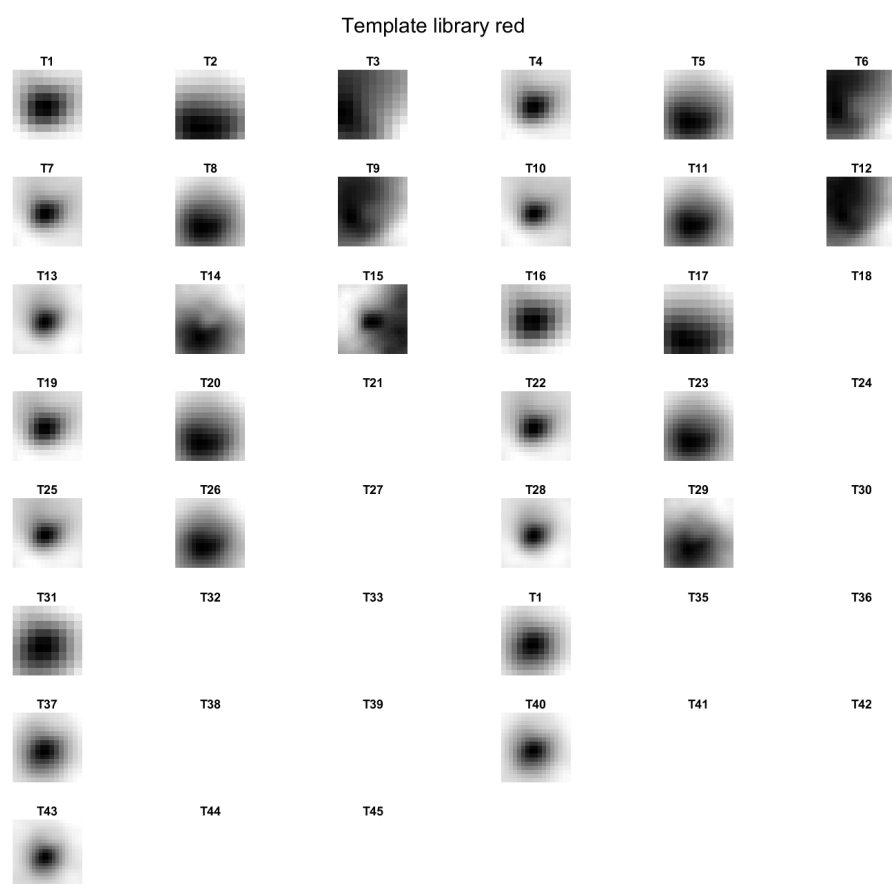


Figure C.1: Template library of the red band generated with eCognition®.

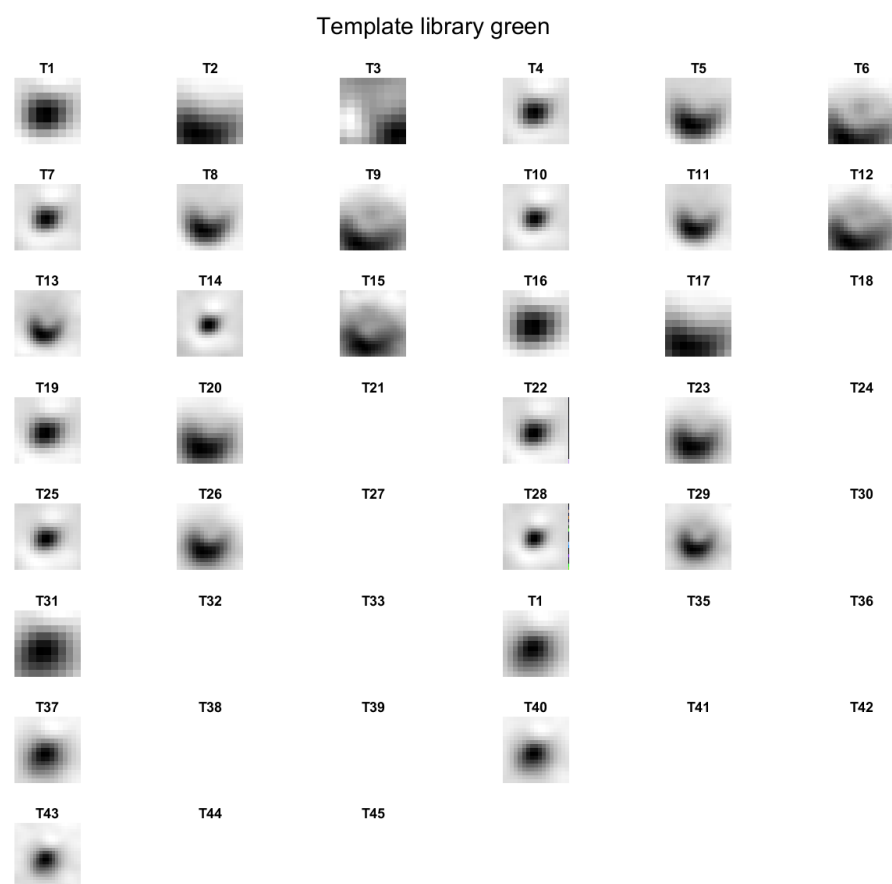


Figure C.2: Template library of the green band generated with eCognition[®].

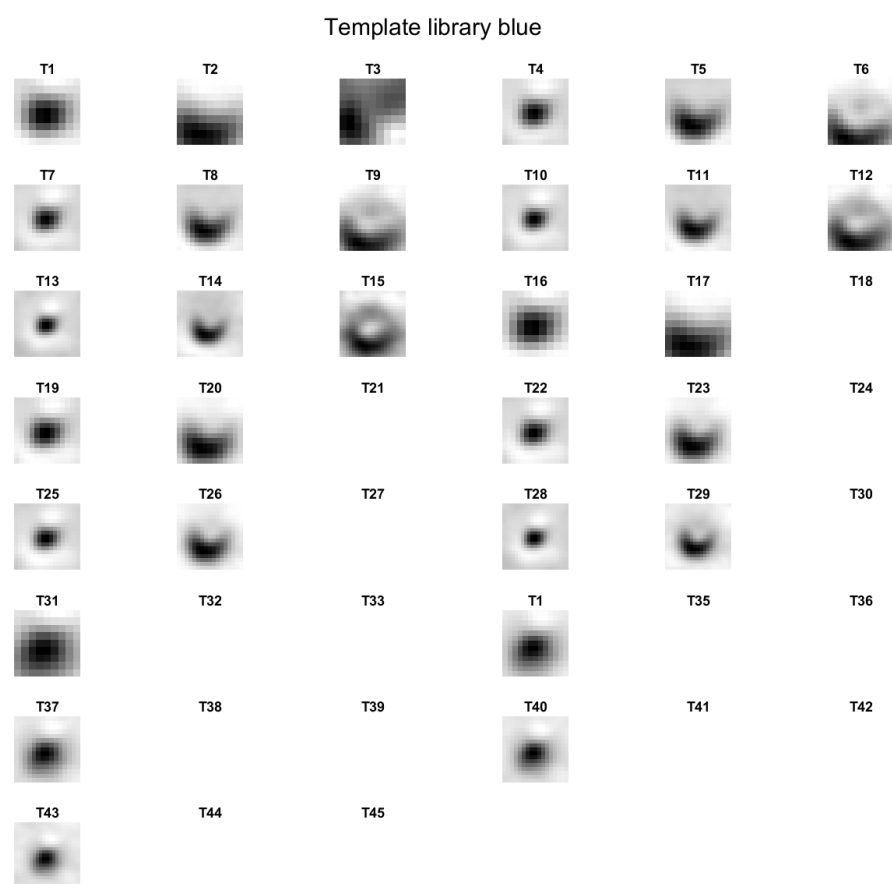


Figure C.3: Template library of the blue band generated with eCognition®.

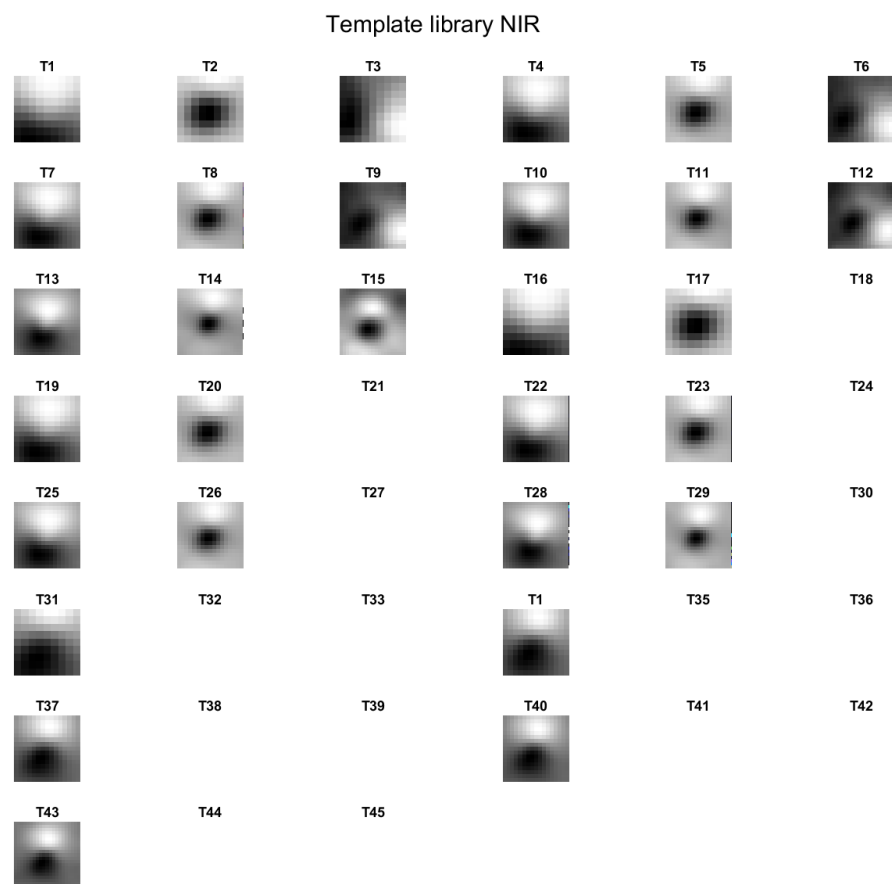


Figure C.4: Template library of the NIR band generated with eCognition®.

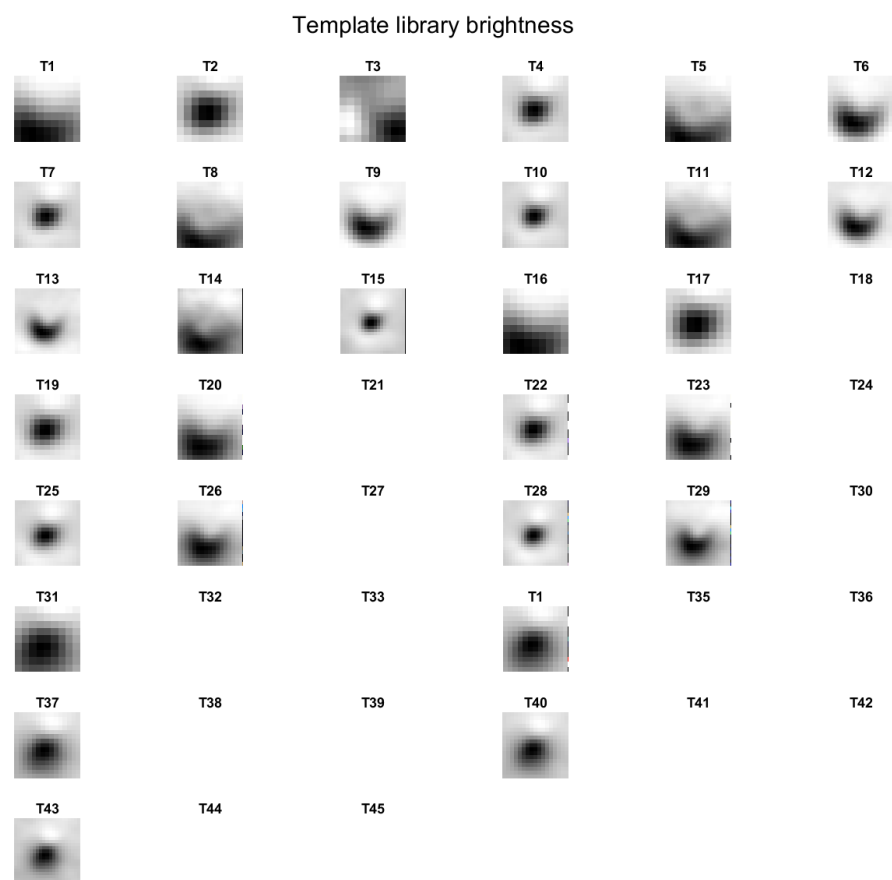


Figure C.5: Template library of the brightness feature generated with eCognition®.

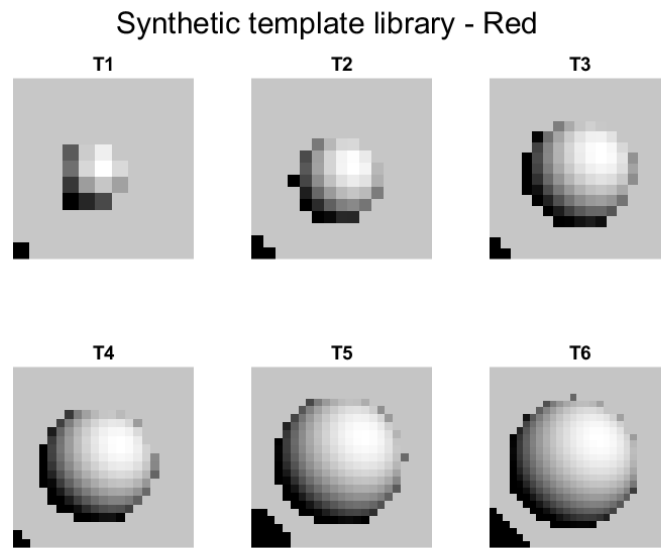


Figure C.6: Template library of the red feature generated with with algorithm 3 in section 3.4.1.

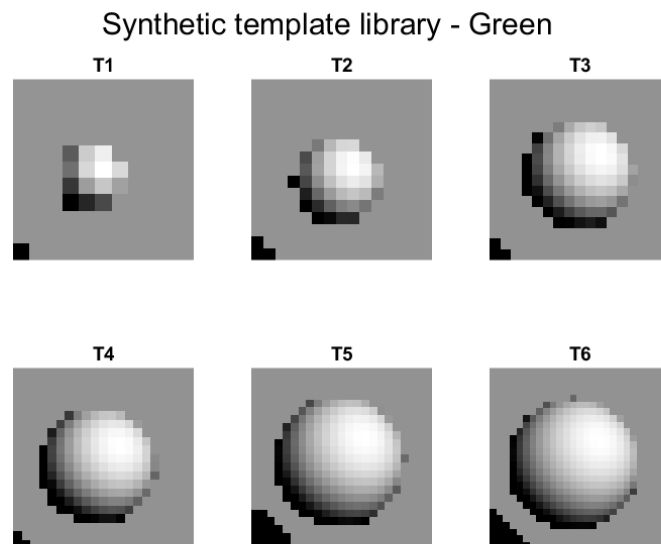


Figure C.7: Template library of the green feature generated with with algorithm 3 in section 3.4.1.

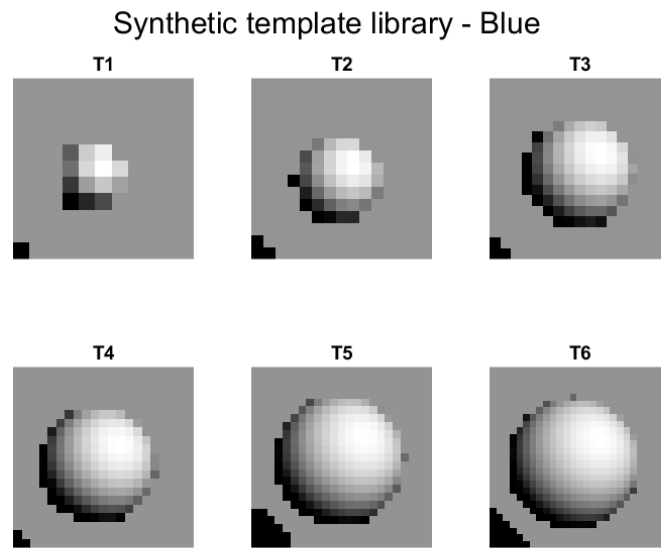


Figure C.8: Template library of the blue feature generated with with algorithm 3 in section 3.4.1.

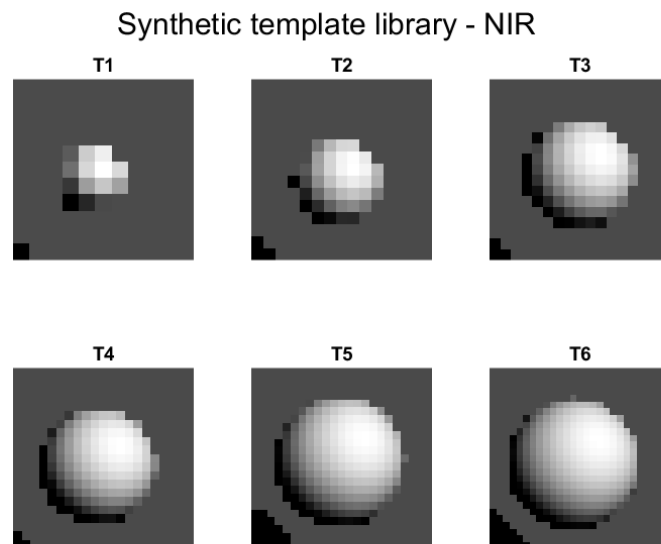


Figure C.9: Template library of the NIR feature generated with with algorithm 3 in section 3.4.1.

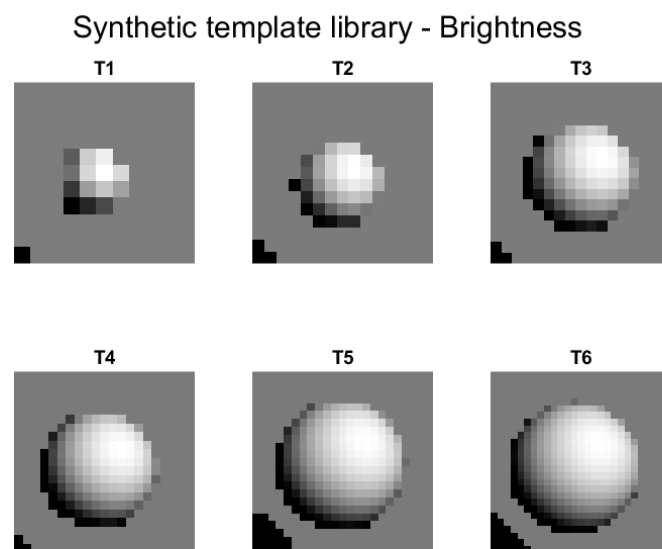


Figure C.10: Template library of the brightness feature generated with with algorithm 3 in section 3.4.1.

C.2 Plantation

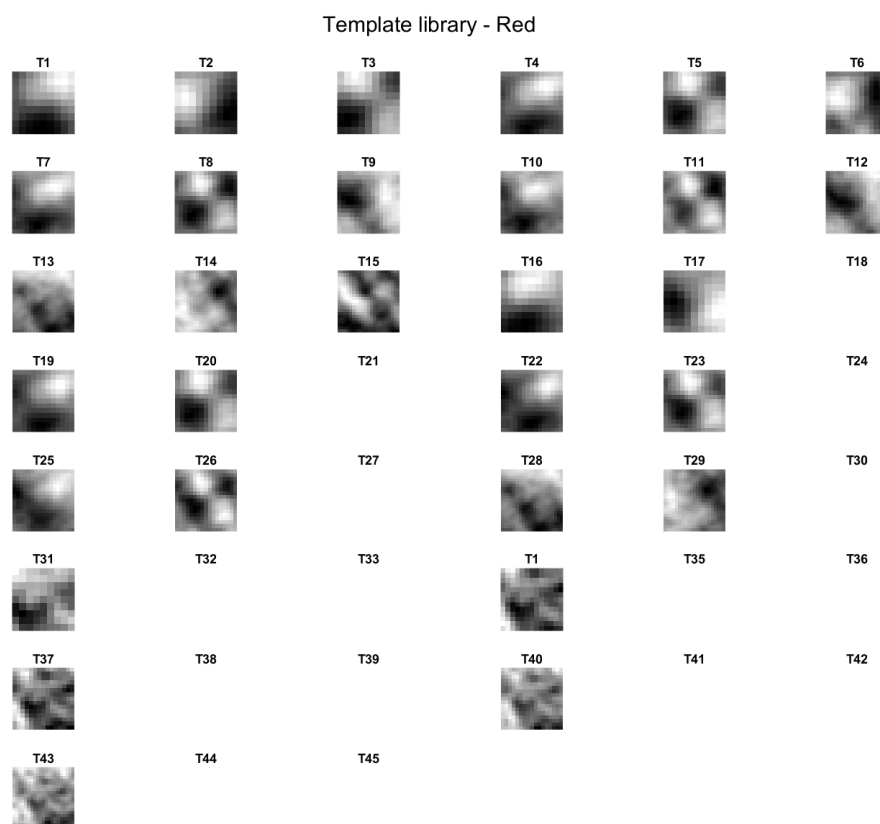


Figure C.11: Template library of the red band generated with eCognition®.

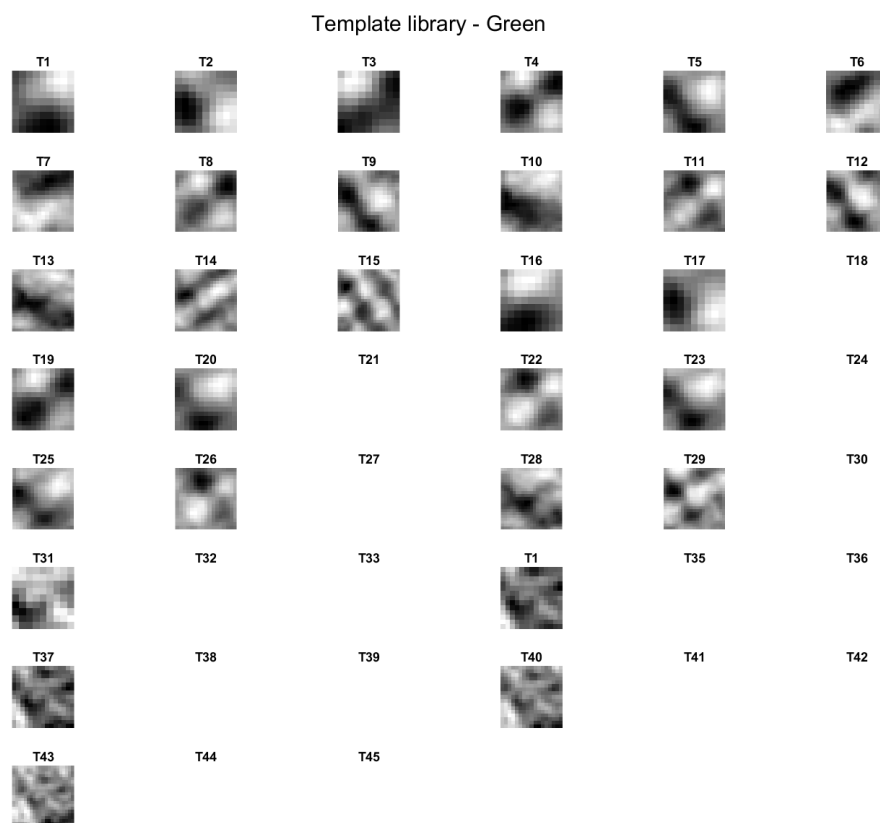


Figure C.12: Template library of the green band generated with eCognition[®].

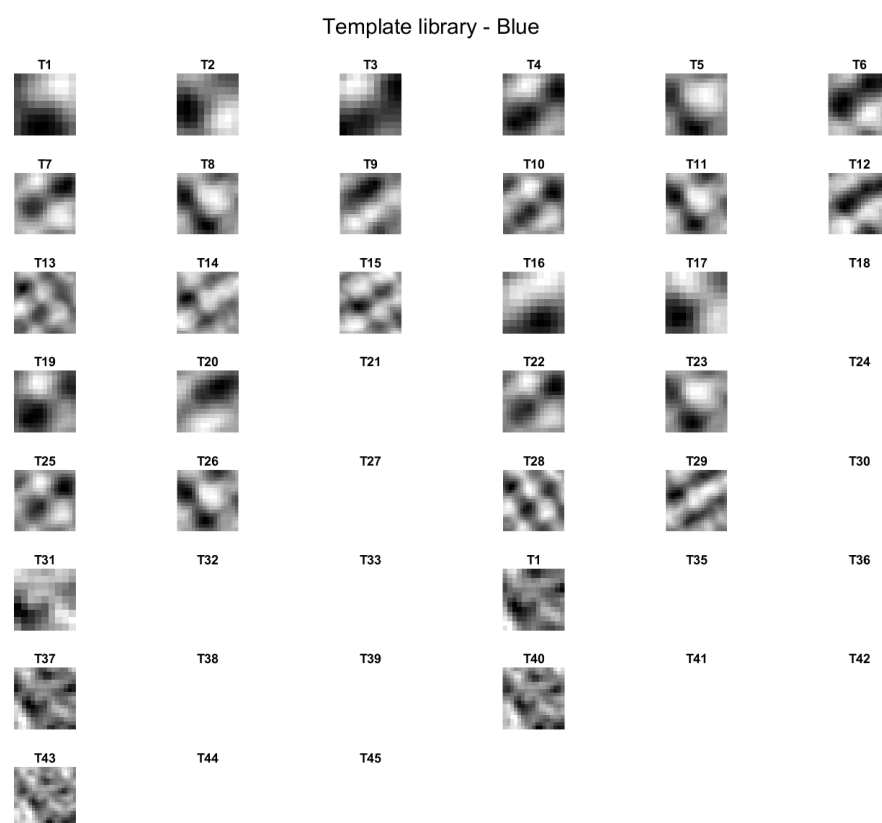


Figure C.13: Template library of the blue band generated with eCognition[®]

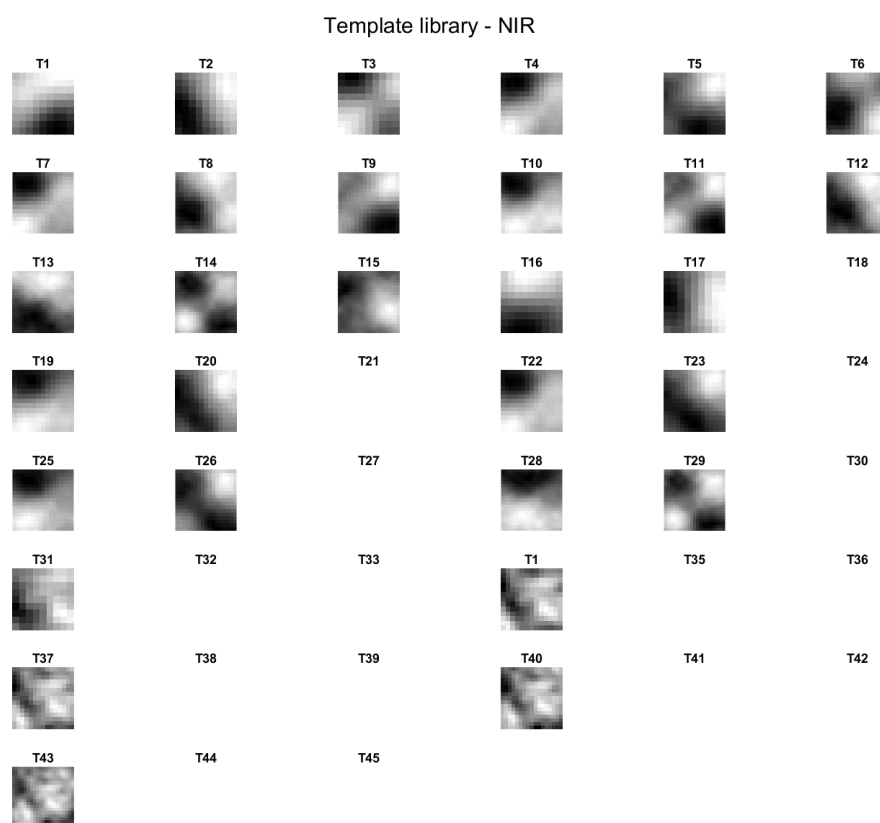


Figure C.14: Template library of the NIR band generated with eCognition[®].

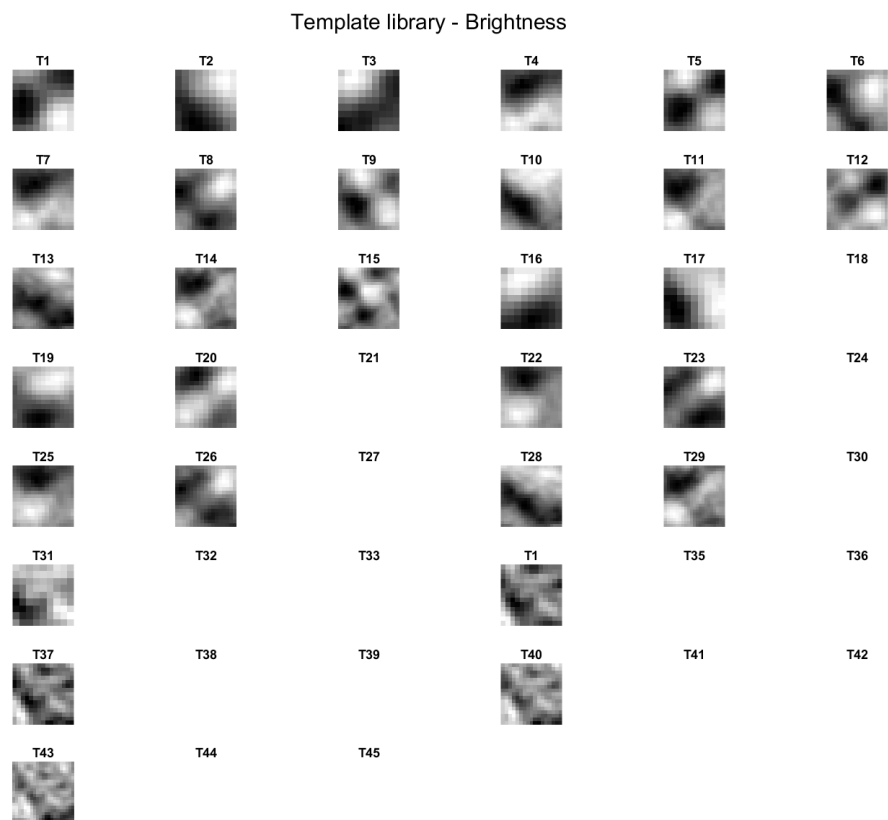


Figure C.15: Template library of the brightness feature generated with eCognition®.

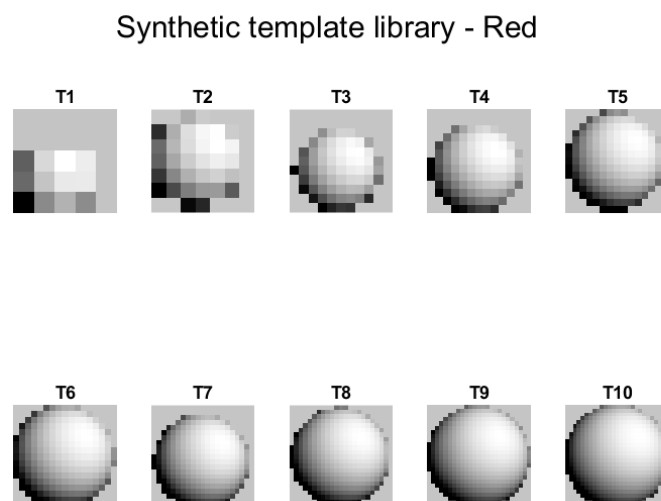


Figure C.16: Template library of the red feature generated with with algorithm 3 in section 3.4.1.

Synthetic template library - Green

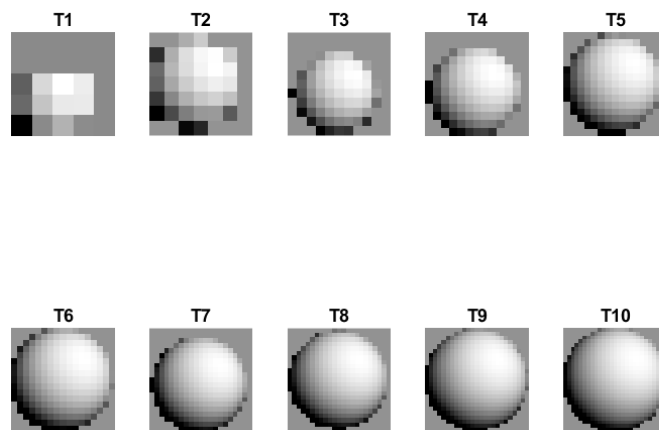


Figure C.17: Template library of the green feature generated with with algorithm 3 in section 3.4.1.

Synthetic template library - Blue

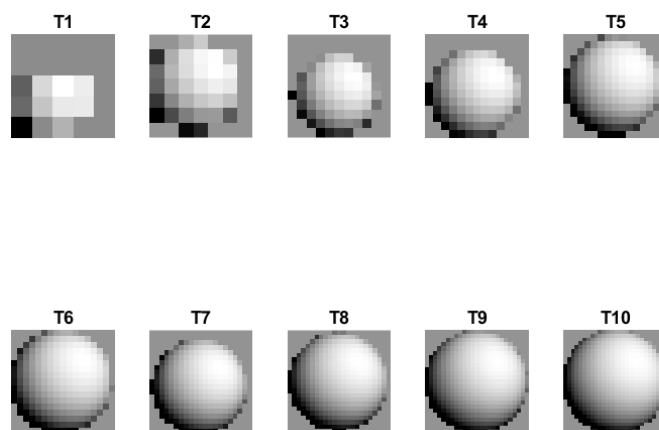


Figure C.18: Template library of the blue feature generated with with algorithm 3 in section 3.4.1.

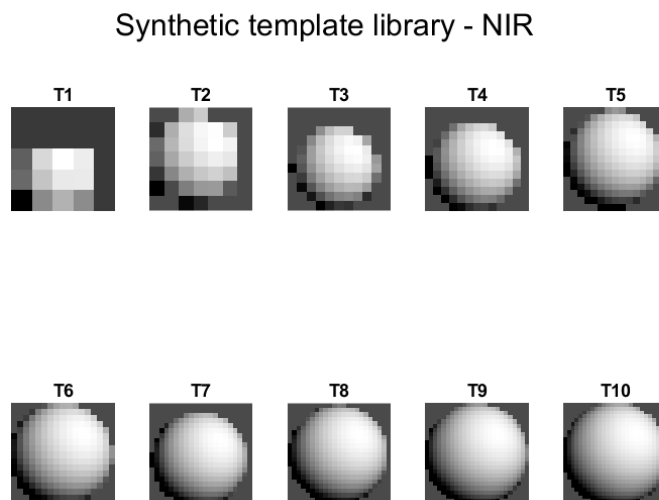


Figure C.19: Template library of the NIR feature generated with with algorithm 3 in section 3.4.1.

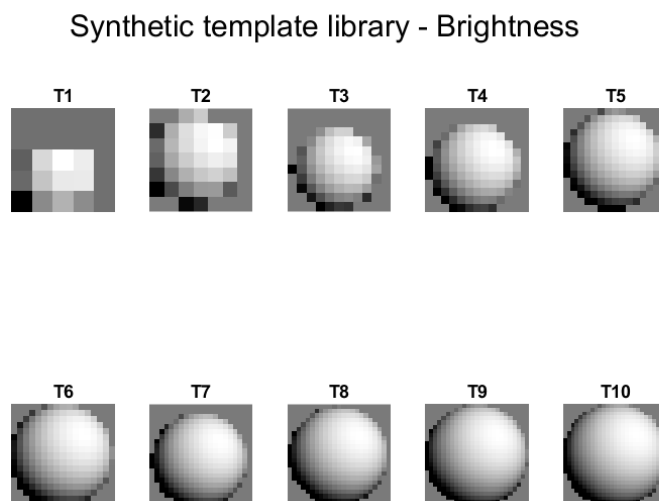


Figure C.20: Template library of the brightness feature generated with with algorithm 3 in section 3.4.1.