



**University of
Zurich**^{UZH}

Developing a Method to Improve SAR Change Detection Under Varying Illumination Angles

GEO 511 Master's Thesis

Author

Lukas Ritter
14-937-817

Supervised by

Dr. Daniel Henke
Dr. Elías Méndez Domínguez

Faculty representative

Prof. Dr. Alexander Damm

30.09.2022

Department of Geography, University of Zurich



Universität
Zürich^{UZH}

Developing a Method to Improve SAR Change Detection Under Varying Illumination Angles

GEO 511 Master's Thesis

Author:

Lukas Ritter
14-937-817

Supervised by:

Dr. Daniel Henke

Dr. Elías Méndez Domínguez

Faculty representative:

Prof. Dr. Alexander Damm

September 30th, 2022

Department of Geography, University of Zurich

Abstract

Change detection is an important field of application for remote sensing. While many studies rely on optical data, Synthetic Aperture Radar brings the advantage of its potential operability being independent from natural illumination and weather conditions. Particularly in time sensitive situations, this can be a great advantage. However, SAR change detection evokes some challenges that may not be seen in other sensor technologies. One of these challenges includes its dependability on a consistent illumination geometry when applying known 2D change detection methods. False alarms increase with growing difference between the illumination angles of a reference and test image. The characteristics of these false alarms have been studied in this thesis and it has been revealed that they predominantly occur as pairs caused by a single object. This phenomenon has been exploited to develop an algorithm which checks elements of change detection on potential pairs and removes them if a certain agreement between the elements of the pairs is met.

The results have shown that the method is able to greatly reduce the occurrence of false alarms, especially in large angular deviations. In smaller angles where current methods perform well, the algorithm retains a majority of the detected changes which are mostly attributable to true changes. Challenges persist in intermediate angles, where false alarms are ample, but pairs are not as pronounced. Nevertheless, the method managed to significantly reduce the false alarm rate in all acquisitions in a circle while retaining the detection rate to a great part.

Table of Contents

1. Introduction	- 1 -
2. Methods	- 4 -
2.1. Study site and data	- 4 -
2.2. Baseline Methodology	- 7 -
2.3. Find Pairs Algorithm	- 9 -
2.3.1. Preprocessing of the Change Mask	- 10 -
2.3.2. Object extraction	- 14 -
2.3.3. Agreement matrix	- 15 -
2.3.4. Normalization, weighting and summation of agreement values	- 20 -
2.3.5. Determination of best match for each object	- 21 -
2.3.6. Selection of pairs for exclusion	- 22 -
2.4. Evaluation and accuracy assessment	- 24 -
2.4.1. Ground truth	- 26 -
2.5. Validation with independent data set	- 27 -
3. Results	- 28 -
3.1. Accuracy assessment of the baseline method	- 28 -
3.1.1. 0118 vs. 0110	- 28 -
3.1.2. 0118 vs. 0111	- 30 -
3.2. Accuracy assessment of change maps after the application of the Find Pairs algorithm	- 31 -
3.2.1. 0118 vs. 0110	- 31 -
3.3. Visual comparison and performance assessment of both methods	- 34 -
3.3.1. 0118 vs. 0110	- 34 -
3.3.2. 0118 vs. 0111	- 37 -
3.4. Weighting and sensitivity to the different agreement criteria	- 40 -
4. Discussion	- 42 -
4.1. Evaluation of the baseline method's results	- 42 -
4.2. Evaluation of Find Pairs' results	- 44 -
4.3. Methodological limitations	- 49 -

4.3.1. Availability of changes for performance evaluation and ground truth _____	- 50 -
4.4. Implications on the applicability of the method developed in this thesis _____	- 51 -
4.5. Outlook _____	- 52 -
4.5.1. Possible improvements and enhancements to the method developed in this thesis - 52 -	
4.5.2. Alternative approaches _____	- 54 -
4.5.3. Potential fields of application _____	- 54 -
5. Conclusion _____	- 56 -
Acknowledgments _____	- 57 -
References _____	- 58 -
Appendix _____	- 63 -
A. Accuracy assessment – baseline _____	- 63 -
B. Accuracy Assessment – Find Pairs _____	- 66 -
C. Accuracy assessment – comparison _____	- 68 -
D. Results of an independent, suboptimal dataset _____	- 71 -
Personal Declaration _____	- 73 -

List of Abbreviations

SAR	Synthetic Aperture Radar
SLC	Single Look Complex
2D	Two-dimensional
3D	Three-dimensional
TomoSAR	Tomographic Synthetic Aperture Radar
DSM	Digital Surface Model
MCC	Matthews Correlation Coefficient
BL	Baseline Method
FP	Find Pairs

1. Introduction

Many geographic applications demand data of a certain standard of spatial and temporal resolution. Remote sensing is a popular data source since it offers a great variety of platforms and sensor technologies, each of which is tailored to meet certain requirements. Observing targets from a distance enables the combination of ambitious needs in both spatial and temporal resolution. A pair of satellites, for example, can revisit the entire globe within a few days while being able to distinguish between objects as small as a few meters. Other platforms like airplanes and drones can offer the coverage of smaller areas with potentially higher resolution and more specialized characteristics for a specific project or research question. Remote sensing can be divided in various sub-disciplines. Commonly, it is differentiated between active and passive remote sensing or between optical and microwave-based remote sensing. Synthetic Aperture Radar (SAR) is the most widespread example for active microwave remote sensing (Woodhouse, 2006, Lillesand et al., 2015).

In today's applications of remote sensing, the other commonly used technology is passive optical and infrared remote sensing. SAR offers the capability to generate high resolution while having the additional advantage of not being reliant on an external energy source, i.e., the presence of sun light. As such, SAR overcomes the main limitation of optical remote sensing by being able to operate both at night and in any condition of sky cover (Patel et al., 2010, Quin et al., 2014).

This is especially useful in situations where data are needed for a specific point in time where deviations caused by meteorological conditions cannot be allowed. Examples for such applications would be rapid mapping or, generally, the observation of catastrophic events where images temporally close before and after an event are needed (Lillesand et al., 2015).

An important field of application of any remote sensing technology is change detection. It encompasses the process of identifying differences in the state of objects or phenomenon at different points in time. Examples of SAR to be used in change detection are, among others, earthquakes, subsidence, glacier and ice cap movements and many more. (Hecheltjen et al., 2014)

Studies of human influences such as infrastructure and urban development are further examples (Mendez Dominguez et al., 2018b).

Especially in situations of risk assessment and in response to catastrophic events, the above-mentioned advantages of SAR come into play, as these application situations of change detection often include adverse weather conditions, where the ability to penetrate clouds is very beneficial. This allows to disable the main constraint for a high temporal resolution which is important in such circumstances.

Conventional SAR applications usually follow a linear acquisition flight path, yielding in different limitations that are typical for radar remote sensing. These limitations mainly include geometrical effects such as foreshortening and shadows. Circular SAR addresses some of these issues as it illuminates a target from different viewing angles. It also offers the possibility to combine the recordings from different angles to a single image which features a greater level of detail. (Lin et al., 2012)

Circular SAR also evokes new challenges, since recordings of a given object from different viewing angles may not perfectly align.

In general, SAR change detection is a field that has mainly relied on a set of methodology that is based on the comparison of texture and intensity values, often also involving thresholding (Gong et al., 2014). Especially the advance of openly available SAR data by the launch of the Sentinel 1 satellites has promoted increasing efforts to enhance methodology to analyze and deploy SAR data in various fields. Nevertheless, some difficulties of SAR based change detection still persist and can be attributed to multiple phenomena that are associated with this technology such as speckle noise and the general complexity of the acquisition geometry system. As a result, many methods with different advantages and disadvantages exist, but none of them can be considered optimal, regardless of whether they are more based on pixel radiometry or on neighborhood statistics (Hachicha and Chaabane, 2014).

In several studies, Mendez Dominguez et al. (2019) have examined various new approaches of SAR change detection, including circularly acquired data. Many of their approaches harvest three-dimensional information, meaning that the acquired data is tomographic SAR data. Furthermore, two-dimensional and three-dimensional change detection methods have been combined to create a change detection method that yields in high accuracies, combining advantages of 2D and 3D data. Targets with a low vertical footprint are assessed under a traditional 2D approach while changes caused by taller targets are detected by also considering their 3D structure. The authors argue that expanding the synthetic aperture in elevation enables to

overcome known SAR limitations such as layover effects and allows to resolve multiple scatterers that are spatially close. As such, it is possible to harvest more of SAR's advantageous characteristics while addressing some of its restrictions. This method forms the underlying basis for this thesis and will be further elaborated in Methods under Section 2.2. It has been shown that even if one applies this newly developed and sophisticated change detection algorithm on circular datasets, the number of false alarms increases with the deviation in angle between two recordings.

Given the fact that there are many potential situations where recordings of a target of interest are available but have been acquired from different viewing angles, it is of great interest to examine the behavior of change detection accuracies under different viewing angles and potential improvements thereof. Under the assumption of different recordings deviating in their viewing angle, which does not necessarily require a circular acquisition but could also be from multiple linear acquisitions that are not parallel, the following research questions have been defined for this thesis:

- How does the quality of SAR change detection vary with different viewing angles?
- How can we improve the performance of existing change detection algorithms with increasing difference in viewing angles between acquisitions?

We intend to assess different approaches of improvement to change detection algorithms and implement the one that is most promising. Ideally, the improvement will increase accuracy levels of the chosen baseline change detection method.

2. Methods

2.1. Study site and data

All data used in this study was sourced by DLR's F-SAR sensor as described by Horn et al. (2008). The sensor offers the ability to record in full polarimetric capability in the X-, C-, S-, L- and P-band. For this project, data recorded in the X-band in VV polarization has been used. The recording occurred in a single-pass interferometric mode with a baseline of 0.81 meters (Mendez Dominguez et al., 2022).

This sensor configuration was used for two acquisitions at two different locations. The first data set consist of both a circular and two linear acquisition flights. For the second data set, a circular flight path was used.

The first data set covers an area of the Allgäu Airport in Memmingen, Bavaria. The area depicted in Figure 1 features different manmade structures, such as buildings of different sizes and structure, parts of the airport's apron with parked narrow body passenger aircraft as well as some natural targets such as trees and some grassland.

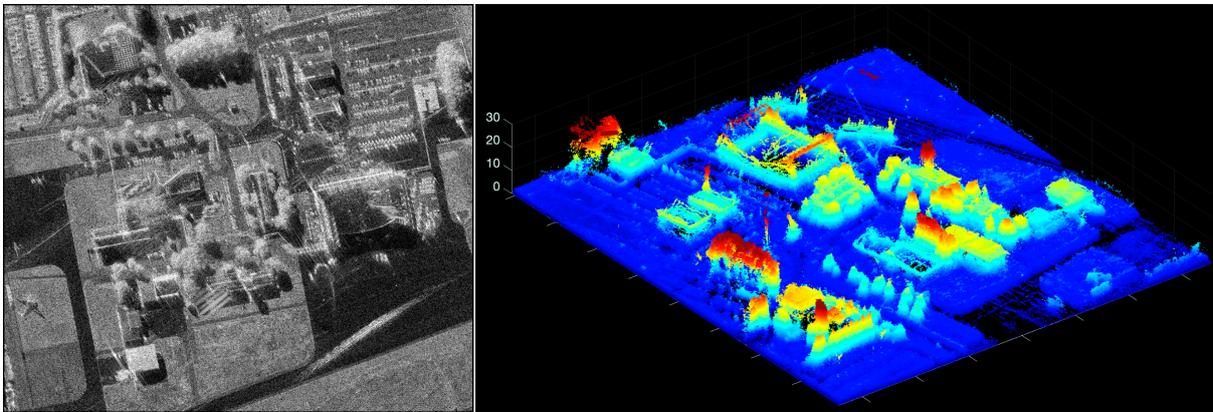


Figure 1: 2D-SAR image from the northern acquisition flight (left) and 3D-SAR DSM (right) of study area I

While all subapertures of the circular acquisition of the Memmingen study area can be considered as being recorded simultaneously, the additional presence of two linear acquisitions enables us to have a time difference between the recordings. All three recordings occurred on the 25th of June 2019 with the two linear flights happening roughly one hour before the circular flight. The time difference between the two linear flights is five minutes.

The circle gives a view of the target area from every direction while the two linear paths are north and south of the study area. The geometry of the flight paths can be seen in Figure 2. One can observe the relatively small area of the spotlighted study site compared to the flight path distance and extent.

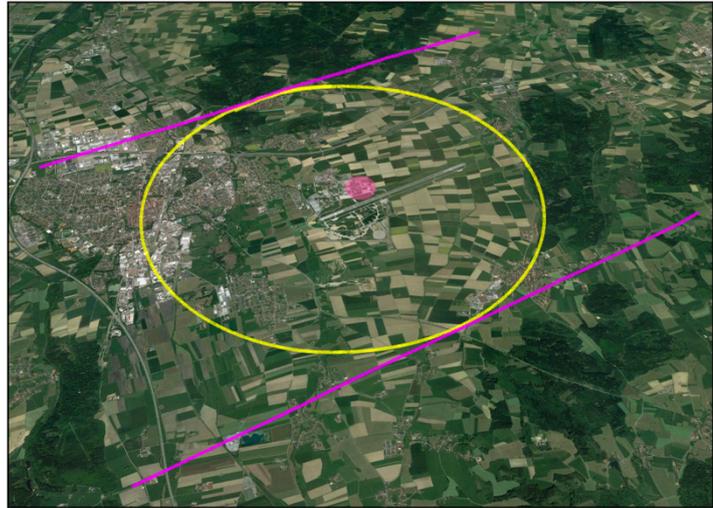


Figure 2: Study area I and acquisition flight paths. Spotlighted area highlighted in magenta. Satellite imagery sourced from Google Earth.

Figure 3 shows all the changes that have happened between the circular and the linear recordings. On the left, we can see an image that is derived from the circular acquisition flight. If we compare it to the image from the linear flight in Figure 1, we can observe the presence of two airplanes on the apron. Other changes mainly consist of cars and other vehicles in the parking lots and some ground service equipment on the apron. The change map displayed in Figure 3 has been created manually and functions as ground truth for this project.

The second study area is located in Friedrichshafen at the shore of Lake Constance. There, subpar flight conditions did not allow to spotlight all subapertures of the circle such that only a part of the recording can be used. The absence of a second acquisition flight means that it can be assumed that there are no changes that could be detected. As a result, this dataset only will be used for comparison and verification purposes.

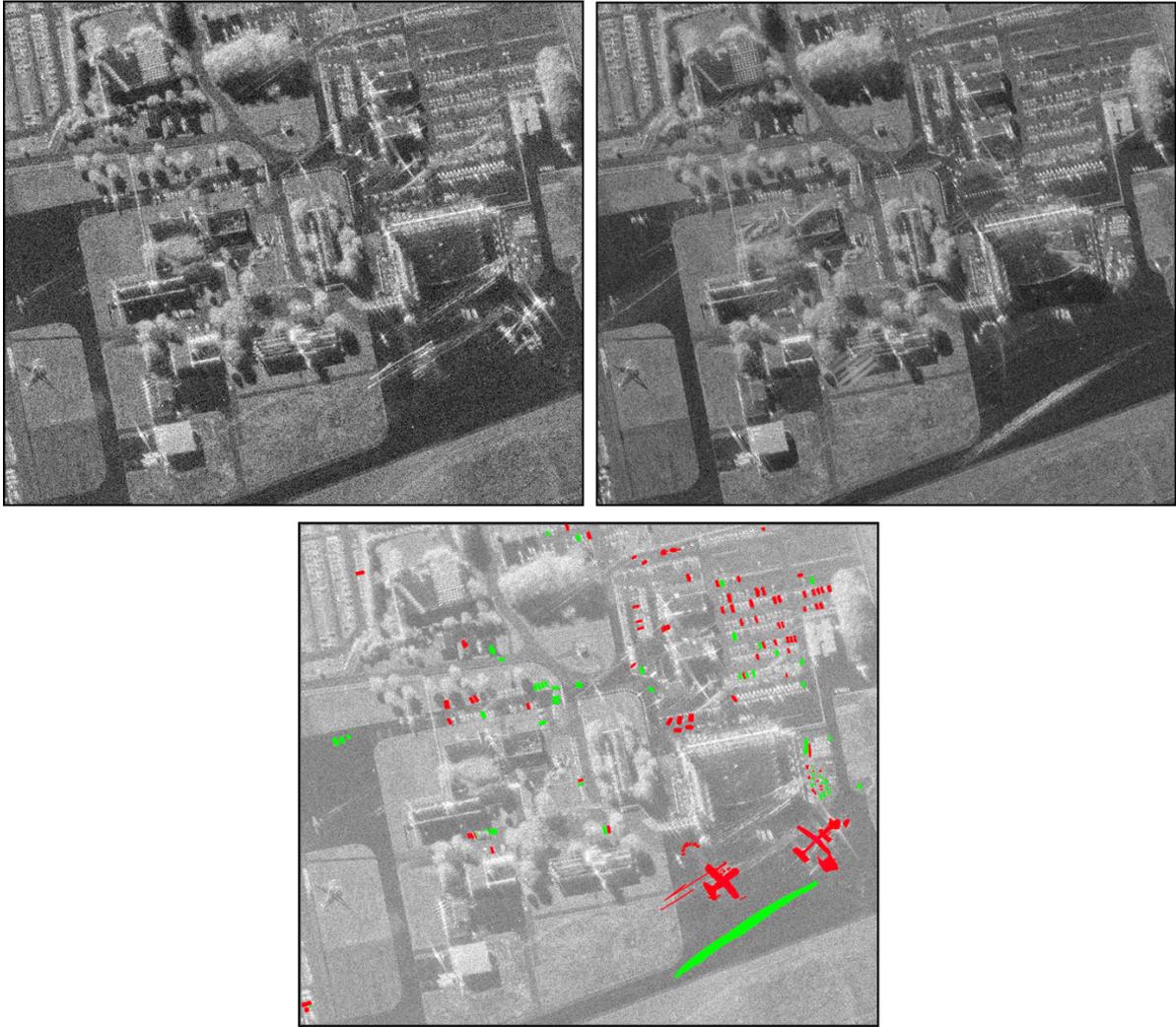


Figure 3: 2D-SAR image from the circular recording (left) and linear recording (right) and with overlaid change map (bottom) of study area I

Table 1 gives an overview of all datasets and their acquisition flight parameters used in this thesis. The circular flight 0118 of Memmingen is subdivided in several subapertures that are obtained by compressing 9000 pulses each and identified by their respective first pulse. There are overlapping subapertures with finer steps of 1000 pulses with first pulses between 1000 and 9000. The rest of the circular data is resolved with non-overlapping steps of 9000 pulses. This implies that between each subaperture step of 9000 pulses, there is an angular difference of $\sim 12.5^\circ$ whereas the finer increment data features a step size of $\sim 1.37^\circ$. Additionally, the subapertures that are closest to their respective linear flight 0110 and 0111 with first pulses 2763 and 122189 have been provided as well. This means that data or an image denoted with “0118-127000” originates from the subaperture with an angular deviation of 177° from the origin.

For the circular data from Friedrichshafen, the naming scheme does not include the first pulse but the subapertures are consecutively numbered.

Table 1: Acquisition flights and datasets used in this thesis

Flight number	First pulse of subaperture	Description / type	Location, date, time
0110	-	northern linear	Memmingen 25-JUN-2019 ~ 11:32
0111	-	southern linear	Memmingen 25-JUN-2019 ~ 11:37
0118	1000 ... 9000 10000 ... 253000	circular	Memmingen 25-JUN-2019 ~ 12:24
0418	1 – 346 usable range: 220 – 240	circular	Friedrichshafen 28-JUN-2019 ~13:41

2.2. Baseline methodology

As already mentioned, the method developed by Mendez Dominguez et al. (2019) builds the basis for this thesis. The motivation behind their approach was to better harvest the full potential of SAR data since 3D information helps to overcome some of its limitations and has so far not been used to address change detection issues. A combination of three steps yields in a method that combines the advantages of 2D and 3D change detection to generate a more reliable change map. A significant improvement in the kappa-coefficient could be shown between the new approach and other established methods. This improvement is attributed to the 3D structure information, a more accurate radar cross section as well as a generally lower sensitivity to known effects like layover and foreshortening. However, the authors state that their method does not account for errors caused by greater differences in the illumination angles between two acquisitions which are believed to be responsible for an increased occurrence of false alarms.

Consequently, the goal of this thesis is to investigate the magnitude of such false alarms in dependence of the difference between illumination angles and find ways to reduce the sensitivity of change detection algorithms to such effects.

While both the 2D as well as the 3D aspect of the method by Mendez Dominguez et al. (2019) have at least visually been considered for the assessment of their susceptibility to differences in illumination, the improvements suggested in this thesis mainly build on the 2D part of the methodology while also optionally considering 3D information. This 2D part which forms the first step of the method is based on other work of a similar set of authors as described in Mendez Dominguez et al. (2018b). The processing chain for the 2D change detection works with a single-look and a multisquint processing mode which are then combined to obtain the difference image. For this thesis, only the single-look mode is used. A difference image is generated by the subtraction of one image from the other. The data is log-transformed beforehand and may be denoised with a block-matching collaborative filter as described by Dabov et al. (2007) and Lebrun (2012) as well. Finally, thresholding using expectation maximization is used to obtain a predefined number of change classes, usually two (change, no change) or three (no change, appearing, leaving targets). This difference image is reported to be more sensitive to sharp targets, as found in urban environments but has inferior performance when looking at environmental features such as trees or grasslands (Mendez Dominguez et al., 2019).

In a second step, after the 2D processing change, the method by Mendez Dominguez et al. (2019) creates a 3D difference image which is preceded by a voxelization of the data. The 3D difference image describes the difference in the occupancy of these voxels. This is combined with information on height difference which is retrieved under the utilization of the Hausdorff distance focusing on differences in vertical position of targets. The combination of the information is done above a certain height threshold using a Conditional Random Field which considers the Euclidian distance and returns the probability for each voxel to be assigned to each possible change class. The background of the Conditional Random Field is described in Sutton and McCallum (2012) as well as Wallach (2004). Its application in SAR change detection has been proposed as well in recently in other projects (Zhang et al., 2021).

For the third processing step, the 2D DI is projected into 3D space using a given ground-height proxy. This builds the base of the 3D change map for the area below the defined height threshold, whereas the result of the second step builds the part

above the threshold. Finally and throughout the processing chain, several mechanisms suppressing false alarms are applied by inhibiting change in cases where voxels are empty.

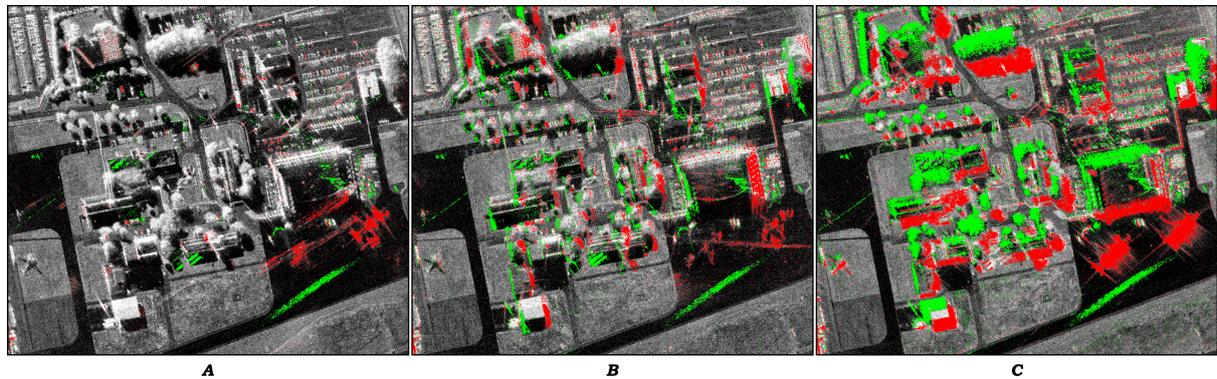


Figure 4: 2D DI and change maps with a difference of illumination angle of $<1^\circ$ (a), $\sim 40^\circ$ (b), $\sim 150^\circ$ (c)

Figure 4 gives a first idea of the relationship between the performance of the change detection and the difference in the illumination angle. While subfigure (a) almost depicts the wanted results of the ground truth in Figure 3, one can observe that already at a difference of 40° , the distribution of false detections is greatly amplified and if the deviation of the angle is increased further, the rate of false alarms vastly exceeds the amount of true positives.

Figure 4 gives one further apparent insight: the pattern of false alarm often is of the way that patches of positive and negative false alarms are mirrored. This can be observed for single and grouped trees as well as building elements. This phenomenon is caused by shadows that are projected in the opposite direction of the illumination and with moving illumination, these shadows move as well, meaning that the disappearing of a shadow creates a positive change and the appearance of the shadow in another pixel neighborhood triggers a negative change.

Although Mendez Dominguez et al.'s combined change detection approach is less susceptible to such shadow effects, the performance decreases significantly with an increased difference in the illumination angle.

2.3. Find Pairs algorithm

The findings in 2.2 gave rise to the idea of a possible improvement. To decrease its sensitivity to the problem of increased false alarms under an increasing difference of the illumination angle, a method has been implemented that will be described in this

chapter. Figure 5 gives a first overview over the sequence of operations of the method developed. The steps involved will be elaborated further in the following subchapters.

The observation that false alarms form pairs suggest developing a way to delete such pairs of false alarms. It gives the advantage that the deletion of true positives is mostly impeded by design since such patterns of change are unlikely to occur in reality. Nevertheless, especially in larger deviations between the angles, a great part of the number of pixels of false alarms can be part of such a pair. Furthermore, the pairs are always spatially related which further decreases the probability of the occurrence of a combination of a positive and negative change that reflects the pattern of a pair. Consequently, the deletion of these pairs should enable the significant reduction of the false alarm rate while preserving the detection rate of the outgoing methodology.

In order to address this approach, an algorithm has been developed, which combines a set of criteria that describe the similarity of objects such that it is possible to determine the likelihood that one object of change forms a pair together with another object of the opposed change class. The working principle of this algorithm will be described in the following subsections.

2.3.1. Preprocessing of the change mask

As a first step, a reference and test SAR SLC image is being loaded into the workspace. The two 2D images are then used to generate a difference image and a change map is being retrieved using the methodology described in 2.2. In this project, a three-class change map is used (no change, incoming, leaving targets). The change map is then separated in two binary change masks (no change, incoming and no change, leaving) which are used for the further processing.

Several preprocessing steps have been considered necessary for the proposed algorithm. As it can be seen in Figure 6, a change mask consists of objects that are large areas of connected pixels as well as smaller objects that are only composed of a few countable pixels and noise that consists of single pixels or very small objects that belong to a larger object but are separated from their corresponding object and also form an object of few pixels.

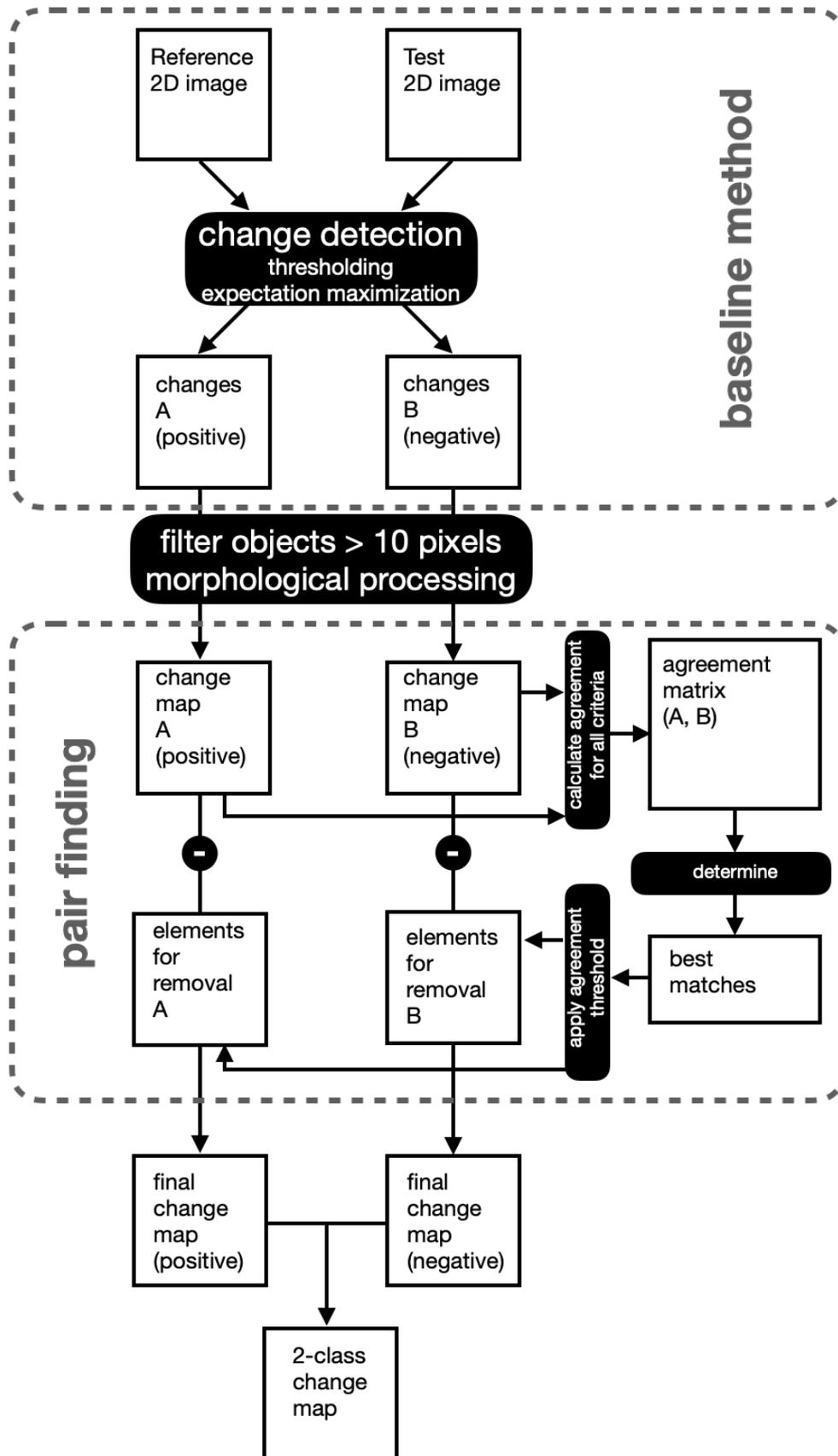


Figure 5: Flowchart of the working order for the Find Pairs algorithm including initial change detection based on the baseline, filtering and morphological preprocessing and the actual pair finding.

In order to find the pairs, it makes sense to have a change mask only consisting of objects that can be compared to other objects. This implies that both noise needs to be removed as well as disconnected parts of single objects should be connected as far as possible.

To achieve this goal, three steps have been conducted. In a first step, all interconnected objects that fall below a defined number of pixels have been removed. For this this thesis, this number has been set to 10, based on an empirical assessment. In order to connect objects and their debris to each other and to smoothen the change mask, morphological image processing has been applied.

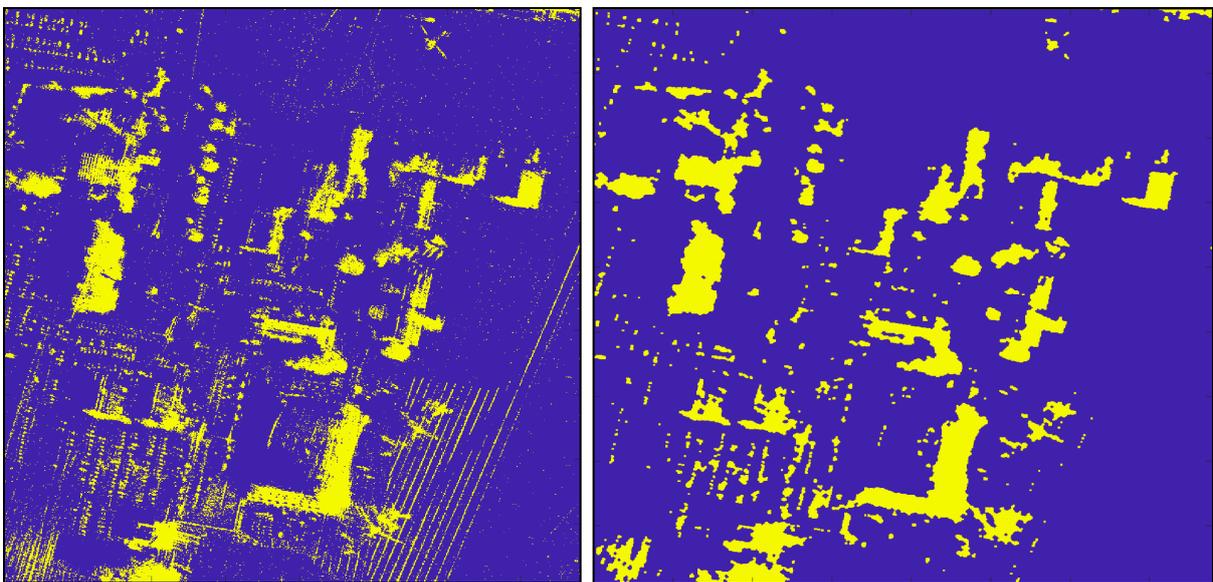


Figure 6: Example of a change map before (left) and after (right) removal of objects and morphological closing followed by opening

In morphological closing, a binary image is first treated with a dilation followed by an erosion. Morphological opening conducts the same process in reverse order. Both operations are conducted using a so-called structure element. In the opening operation, foreground structure smaller than the structure element will be removed, while the same applies for background elements in the closing. The result of the morphological operations is dependent of the structure element and patterns of the structure element are always visible in the result. (Said et al., 2016) As a welcome side effect, image closing tends to connect foreground elements that are spatially close.

For this thesis, the aim was to minimize the morphological impact of the structure element on the output change mask while retaining the intended smoothing and denoising effect of the operation. Since the input change mask is expected to feature

both edge as well as rounded elements, there is no best shape for a structure element. Using disk-shaped, rectangular or diamond-shaped structure elements never delivered the desired results, since both in opening and closing, the structure element is always very apparent in the result, even when using structure elements of only a few pixels in size.

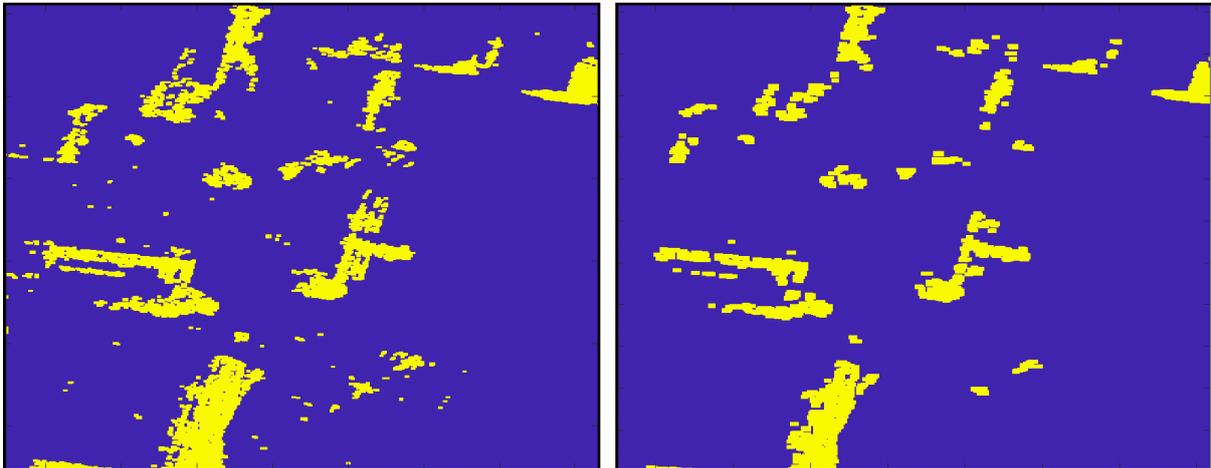


Figure 7: Image closing performed with a disk-shaped (left) and square (right) structure element. The impact of the structure element's shape on the resulting mask is visible.

Figure 7 exemplifies this problem. Not only are the effects of the structure element's shape very pronounced, but also is the smoothing and denoising effect not meeting the desired result. Especially in the left example, many objects of a few pixels in size remain while most larger objects still feature many subparts and small spin-offs.

To overcome this issue, a big number of different variations of structure element shapes and sizes, as well as different sequences of the morphological operations finally led to the most satisfying result, which is depicted in the right part of Figure 6. In this configuration, the objects still reflect their counterparts of the original change mask, with minimized signature of a structure element, while the noise could be greatly reduced and objects are mostly distinct, connected entities.

This result was achieved by first conducting a closing operation with a disk-shaped structure element with a radius of 2 pixels. In a second step, an opening operation using a 3-by-3 square shaped structure element has been applied. Here, the principle that morphological opening and closing are not the inverse of each other despite being the reverse order of the same operations, has been used. This implies that the original cannot be restored through this sequence of operations, offering a meaningful enhancement of the image instead (Richards, 2022).

The above-described sequence of operations has been used throughout the thesis with the option to change size and shape of the structure elements within the code's

parameters. The operations were implemented using MATLAB's built-in functions and toolboxes (MATLAB, 2021).

The work steps described in this section are applied to the two change masks that have been derived from the initial change map. This results in two modified change masks analogous to the example displayed in Figure 6.

2.3.2. Object extraction

After the steps taken in Section 2.3.1, the remaining elements that are formed of connected pixels need to be registered as individual units. For this, an object extraction algorithm based on Klette and Rosenfeld (2004) has been applied. The algorithm takes the binary mask as an input and treats every set of connected pixels, i.e., all neighboring cells of a matrix containing the same value as a unit. A new image is generated, assigning a unique integer value to all pixels of a connected object. Additionally, the function returns a list containing some geometrical characteristics of each object such as the area, circumference and the centroid. If a background value image is given as an optional input to the function, statistical measures for the background values of each object are given as well. Both the change image as well as the characteristics list are generated for both modified change masks originating from Section 2.3.1.

For the sake of computational efficiency as well as keeping a better overview of the entire process, only objects with a size of at least 100 pixels were taken into the algorithm. This size threshold is set as an addition to the removal of objects below a pixel size of 10 before the morphological processing. Limiting the number of pixels for an object has been done in two steps since the image closing and opening brings a significant smoothing and denoising effect, which is why a relatively small threshold is set initially. After this preprocessing step, the additional threshold is set since those objects below 100 pixels in size are considered not to be noise, such that these objects are excluded from the pair finding procedure, since it is assumed that the finding of pairs would not perform well for such small objects. Furthermore, objects of such a small size may be not as stable, also leading to outliers in the object extraction process which could interfere with the pair finding process. Examples of

such small changes in this project are mainly cars in the parking lot. Since they occur in such a great number and are all very similar in size and shape, it is expected that the pair finding would cancel most of these changes, although they are true positives. The risk of cancelling those changes is further amplified by the fact that these changes of cars moving in and out may be spatially very close, which is an important factor for the determination of the agreement, what is further elaborated in Section 2.3.3.2. Therefore, it has been decided to exclude such small items from the pair cancelling procedure but to feed them back into the change mask after the procedure, since they are considered to be true changes and mostly not noise.

As a result, we have a set of objects A, which are objects that are present in the reference image but not in the test image and a set of objects B featuring objects present in the test image but absent in the reference image. The object extraction algorithm allows to identify and distinguish between those objects that were just clusters of connected pixels beforehand by providing a matrix featuring individual values for each object as well as a collection of characteristics for all objects in both sets A and B.

2.3.3. Agreement matrix

The core part of the method developed in this thesis is formed by a matrix describing the agreement between each object in set A to each object in set B. Therefore, the matrix is of the following size:

$$size_{agreement} = n_{objs A} \times n_{objs B}. \quad (I)$$

An agreement value for each potential combination of an object in A and an object in B is calculated and stored in the matrix. Subsequently, the information in this agreement matrix can be used to assess which combination of objects are likely to form a pair that is then excluded from the change mask.

Each potential pair of objects is assessed on their similarity considering a set of different criteria. Each criterium returns a certain agreement value ac and the final agreement value is formed by the weighted sum of all k individual agreement values. As such, the agreement a between two objects A_i and B_j is given by the following formula:

$$\begin{aligned}
a(A_i, B_j) &= \sum_1^k ac_k * w_k, \\
i &\in \{1 \dots n_{objSA}\} \\
j &\in \{1 \dots n_{objSB}\}
\end{aligned}
\tag{II}$$

where w_k forms the weight for the respective criterium.

In the following subsections, each criterium and its calculation are explained. The agreement value describes the dissimilarity between two objects which implies that the lower the value, the better the agreement.

2.3.3.1. Agreement in area

Pairs of false alarms that are caused by the same object through the difference in the illumination angle are believed to be related in several parameters. The selection of the criteria is based on this assumption and one of the most apparent features that can be expected to be closely related is the size of the two objects. In this case, the size or area of an object can be approximated best by the number of pixels.

The agreement value for the area was formed by the quotient between the difference of the number of pixels and the area of object A:

$$a_{area}(A_i, B_j) = \frac{|area_{A_i} - area_{B_j}|}{area_{A_i}}
\tag{III}$$

2.3.3.2. Agreement in distance

A second indicator for the probability of a pair to belong together is their distance. The distance is depending in the amplitude of the angle difference and of the object size. Below a certain threshold, distance values are in a realistic range and depending on the width and height of a scatterer, distance values can be considered as being equally likely to be an indicator for a possible relation. Above this threshold, however, a relation becomes increasingly improbable.

Due to this modality, it has been decided not to use a linear increase in the agreement value with increasing distance. Instead, a function was formed that reflects the above-described case. Below a defined threshold t , an exponential function ascends merely until close to the threshold followed by a steep increase, making a relation unlikely with increasing distance. After a certain distance, where the agreement

values have reached unrealistic values, the function flattens again since a further increase in distance does not make the already unrealistic relationship significantly more unlikely. The agreement function has been defined as follows:

$$a_{\text{distance}}(A_i, B_j) = \begin{cases} 0.2 * 1.2^{0.12 * \text{dist}_{A(i)-B(j)}}, & \text{for } 0 < \text{dist}_{A(i)-B(j)} \leq t \\ \sqrt{\text{dist}} + 31, & \text{for } \text{dist}_{A(i)-B(j)} > t \end{cases} \quad (\text{IV})$$

where

$$\text{dist}_{A(i)-B(j)} = \sqrt{(x_{A(i)} - x_{B(j)})^2 + (y_{A(i)} - y_{B(j)})^2}. \quad (\text{V})$$

Equation (V) describes that the distance between two objects is calculated as the length of the vector between the two centroids of the objects $A(i)$ and $B(j)$.

Based on the size of buildings and other targets in the scene as well as the extent of the images, t has been set at 230. In doing so, the distance-agreement curve assumes unrealistically high values if the distance of two objects is significantly greater than 200 pixels. This behavior can be observed in Figure 8. Both the function parameters as well as the parameter t have been tuned for the F-SAR recording in Memmingen (study area I) and are expected to work in similar environments with similar

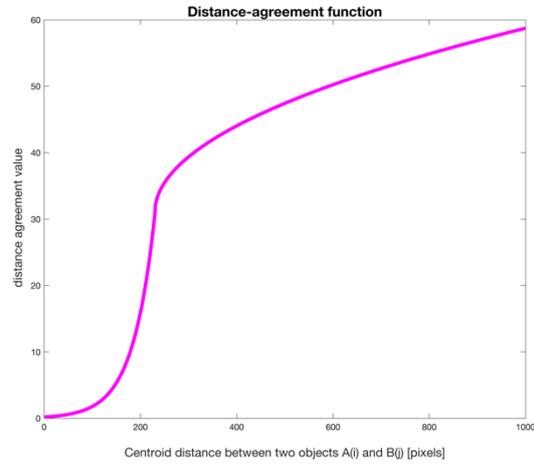


Figure 8: Function describing the likelihood of a relation between two objects depending on their distance. $t = 230$ as described in equation (IV)

setup of sensor and flight paths but might need adjustments if applied to other datasets.

2.3.3.3. Agreement in height

The occurrence of pairs of false alarms has been observed to be caused by objects of a certain size and, especially, height. This can be attributed to the geometric characteristic of the illumination and the resulting movement of a shadow around an object. This implies that between two false alarms forming a pair, the presence of a taller object is to be expected, offering the opportunity to include one further measure to the evaluation of the agreement of two objects.

In order to assess this feature, a buffer zone is created between the centroids of two objects. Due to the lack of a feasible function, this buffer zone has been implemented by a separate function. The size of the buffer is defined by the bounding rectangle of each object. A visualized example of the determination of the height agreement can be seen in Figure 9. Subfigure A shows the investigated pair of objects. This pattern of false alarms is created by a group of tall trees. The buffer zone is displayed in subfigure B. The height matrix is visualized in C. It is being retrieved by denoting the highest occupied voxel in the 3D-SAR image for each ground pixel. Therefore, it is comparable to the SAR DSM in Figure 1.

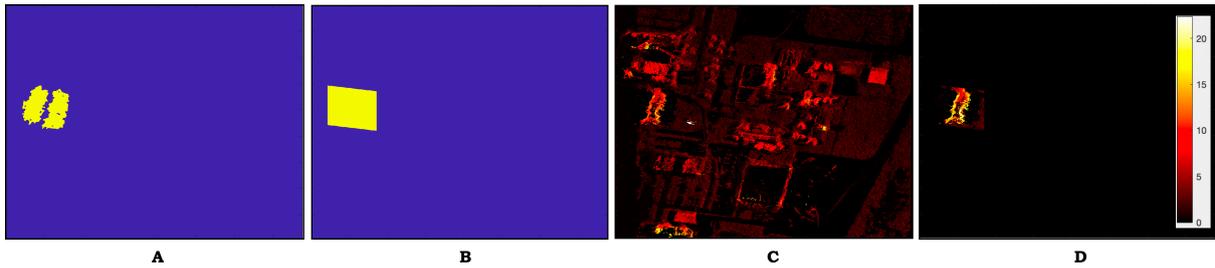


Figure 9: Process of height agreement evaluation. Pair of objects retrieved from both change masks (A), computed buffer between objects (B), height matrix (C) and considered height masked with buffer in D.

Contrary to that example, the height info used here only exploits the 3D SAR data that is available from the reference and test image. For the image in Figure 1, data of all subapertures of the circular acquisitions have been summed up. Here, we intentionally only used data of a single subaperture in order to simulate the limited availability of recordings in an application. Nevertheless, one can see that this limited dataset is capable to display the rough height structures as one would expect it for the scene given the distribution of tall vegetation and buildings. The group of trees that presumably causes the pair of false alarms is well defined in the masked area, featuring patterns of height values of 5 – 12 meters and even some areas with 15 – 20 meters.

The calculation of the height agreement value is determined by the percentage of masked pixels that are above a certain threshold. Since a system of low values for a higher agreement is used, the value is inverted by subtracting the retrieved percentage from 1:

$$a_{height}(A_i, B_j) = 1 - \frac{n_{pixels > t_{height}}}{n_{pixels}} \quad (VI)$$

The usage of this criterion is intended to be optional as it requires the availability of 3D data which is not always given.

2.3.3.4. Further agreements in geometrical properties

Similar to the approach in Section 2.3.3.1, additional geometrical features have been taken into the agreement assessment to determine the similarity of objects. These include the object's circumference, their compactness as well as the number of changes in direction of the contour line. All values are computed by and retrieved from the object extraction algorithm introduced in Section 2.3.2. The features enable to further mathematically describe the similarity of geometrical objects. Besides area, circumference is a size parameter. The signal of the two, as for all of these geometrical features, is not independent but nevertheless, a shape with the same area may have a very different circumference, which is also considered in their compactness, which puts the two into a relation. The number of changes in direction can give an additional indication about the regularity or complexity of a given shape. All of this is expected to be similar for objects that shall form a pair.

The calculation of the agreement value for the circumference and the number of changes in direction follow the same principle as for the area explained in Section 2.3.3.1 and equation (III): the absolute value of the difference is set into relation to the value of object A.

For the compactness c , the quotient of the shape with the larger value to the one with the smaller value is built:

$$a_{\text{compactness}}(A_i, B_j) = \begin{cases} c_{A(i)}/c_{B(j)}, & \text{if } c_{A(i)} > c_{B(j)} \\ c_{B(j)}/c_{A(i)}, & \text{if } c_{A(i)} \leq c_{B(j)} \end{cases} \quad (\text{VII})$$

2.3.3.5. Agreement in contrast of SAR image

Besides investigating objects on their geometrical similarity, statistics based on the pixel intensity have been included into the consideration as well. These statistics are not linked to the change mask itself but to the radiometry of the 2D SAR image for the area masked by an object. The standard deviation of the pixel values give a good indication of the contrast of an image (Moulden et al., 1990), especially here since

applied only to a local area within an image. While the average gray value of a target in a SAR image may differ between two different acquisitions due to the difference in illumination, we can expect that the contrast, i.e., the standard deviation offers some more consistency.

The agreement in the standard deviation is calculated as the absolute value of the difference of the standard deviation between the two objects. For object $A(i)$, the standard deviation is retrieved from the reference image and for $B(j)$ from the test image, as it can be seen in equation (VIII):

$$a_{\text{standard deviation}}(A_i, B_j) = |sd(ref_{A(i)}) - sd(test_{B(j)})| \quad (\text{VIII})$$

2.3.3.6. Agreement in the median gray value

Despite its presumed to be limited informative value, the median gray value forms another criterion related to the pixel-based information. Similar to the procedure in 2.3.3.5, the operation retrieves the values of the respective SAR image for the masked area of each object and calculates the median gray value. The agreement is given as the absolute difference between the two medians:

$$a_{\text{median}}(A_i, B_j) = |median(ref_{A(i)}) - median(test_{B(j)})| \quad (\text{IX})$$

2.3.4. Normalization, weighting and summation of agreement values

All agreement values are computed in a nested loop which iterates over all objects in A and B such that an agreement matrix exists for each criterion in the same size as described in equation (I). Following the calculation, all agreement matrices are normalized by their maximum such that the value range for all matrices is [0 ... 1]. As defined in equation (II), the final agreement matrix is the weighted sum of all individual agreement matrices. The weights are determined empirically by trying to optimize the results of the find pairs algorithm. A training set containing change masks from 5 different subapertures has been chosen and several objects that are obvious pairs of false alarms as well as objects of true positives have been observed

during the manual maximization with the goal to find and exclude the largest possible number of false alarms while retaining all true positives. An overview of the weights used for each result is given in Result under Section 3.3.

2.3.5. Determination of best match for each object

In order to determine the pairs for exclusion of the change mask, the best match for each object is to be found. The best match for each object in one set, e.g. $A(i)$, is defined as the object in the other set – $B(j)$ – in which the combination of the two feature the lowest occurring minimum value in the agreement matrix:

$$\text{bestMatch}(A_i) = B_j \mid a(A_i, B_j) = \min(a(A_i, :)) \quad (\text{X})$$

Analogous to this, the best match for an object in B is defined as follows:

$$\text{bestMatch}(B_j) = A_i \mid a(A_i, B_j) = \min(a(:, B_j)) \quad (\text{XI})$$

Since this simple definition leads to cases where objects of one set are assigned as the best match to multiple objects in the other set, the resulting list of pairs to exclude would not be unambiguous. To address this issue, a duplicate handling function has been implemented that checks for any ambiguity and recursively assigns such elements of one set to the element of the other set such that best agreement value possible is achieved. For the other elements in the other set that were initially also assigned to this element, the second, third, ..., n^{th} best match is determined. The function calls itself recursively for each ambiguity until all ambiguities for the current element are resolved and the distribution of assignments is optimal. The working principle of this duplicate handling function is exemplified in the following pseudo code block:

```

bestMatch(Ai) = Bj | a(Ai, Bj) = min(agreement(Ai))    % find best match
bestMatchesA(i) = bestMatch(Ai)
if bestMatch(Ai) in bestMatchesA:                          % check for ambiguity
    n = 2
    bestMatch(Ai) = find_nth_bestMatch(bestMatch(Ai), bestMatchesA, n)

find_nth_bestMatch(bestMatch(Ai), bestMatchesA, n):
    bestMatch(Ai) → Ax | a(Ax, Bj) = min(agreement(Bj))    (XII)
    bestMatchesA(x) = bestMatch(Ax)                            % reassign bestMatch
    bestMatch(Ai) = nth best match for Ai                    % find nth best match
    bestMatchesA(i) = bestMatch(Ai)

if bestMatch(Ai) in bestMatches(A): % recursive call if again ambiguous
    n = n + 1
    bestMatch(Ai) = find_nth_bestMatch(bestMatch(Ai), bestMatchesA, n)

```

The set with the larger number of elements determines which list of best matches is used for the selection of pairs for exclusion.

2.3.6. Selection of pairs for exclusion

In Section 2.3.5 , two lists of pairs of presumptive false alarms have been created. In theory, it would be possible to remove all these pairs from the change masks. This, however, would not lead to a meaningful result, since the lists of best matches are only limited by the availability of potential elements for the formation of such pairs. It would be wrong to assume that all elements are false alarms. In order to avoid the exclusion of correctly identified changes, a threshold has been applied, limiting the exclusion to pairs that fall below a certain agreement value. For this, the distribution of agreement values has been examined using the histogram as well as value samples of selected pairs.

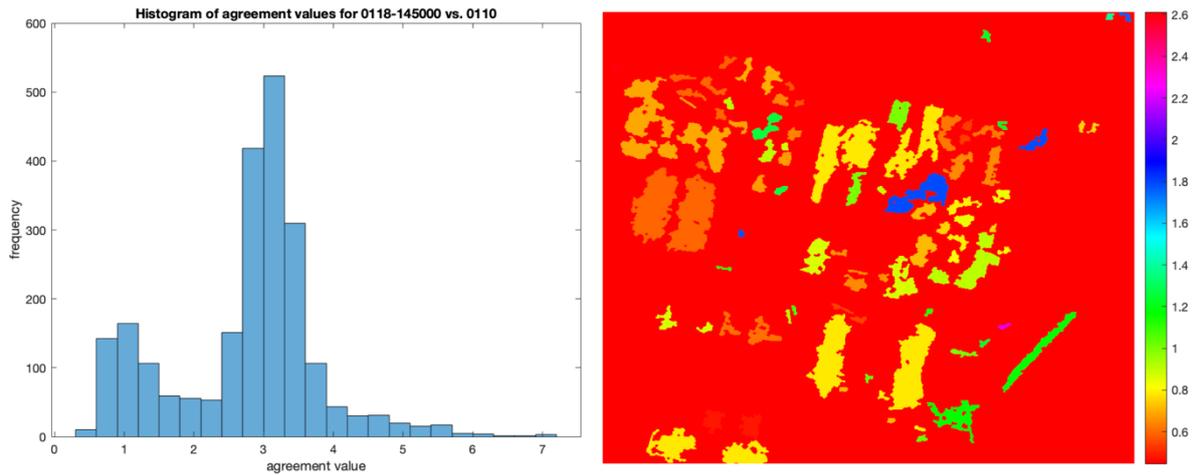


Figure 10: Distribution of agreement values for the 200° circular image against the northern circular acquisition (left). Agreement values visualized for individual elements of change (right) Number of elements in a $n(a) = 54$ and $n(b) = 42$

Looking at Figure 10 might suggest setting the agreement threshold around the first local minimum at a value around 2. However, knowing that a change element originating from an airplane that has moved between the acquisitions features an agreement value of around 1.16 for its best match, we know that the threshold certainly has to be below this value. Looking at the histogram again, a value of 0.8 could deliver promising results. Furthermore, the agreement values of more and less easily recognizable pairs may be used to get additional insight on the distribution and can give an indication for a potential threshold value.

For the present thesis, this empirical approach has been used to set a threshold which was then used for an entire dataset of all parts of a circle versus a linear acquisition.

Once a threshold has been set, all pairs that fall within the value range are used to create removal mask which are then subtracted from their respective change mask that resulted from the preprocessing steps conducted in Section 2.3.1.

An example of the combined change mask (positive and negative changes summed resulting in a 2-class change mask) before and after applying all steps described in 2.3 and subchapters can be observed in Figure 11.

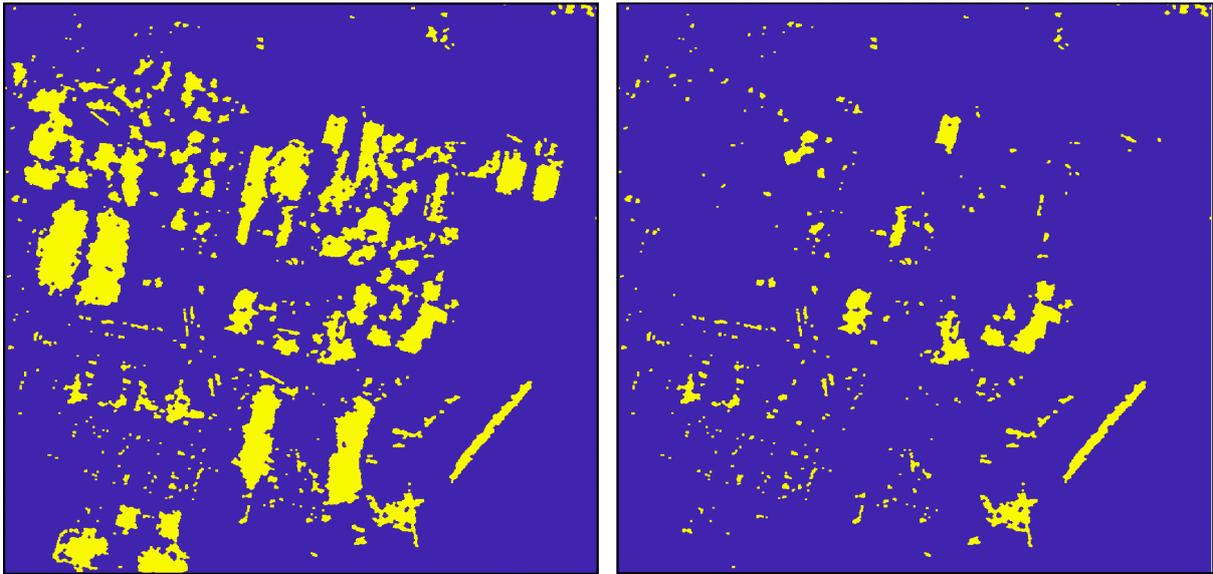


Figure 11: Combined change map before (left) and after (right) performing the find pairs algorithm. 0118-145000 as reference, 0110 as test input image.

2.4. Evaluation and accuracy assessment

Since change detection essentially is a classification problem, the known approach of accuracy assessment as described by Lillesand et al. (2015) will be applied to determine the performance of both the baseline methodology as well as to determine the magnitude of potential improvement achieved through the approaches developed in this thesis. The most important terms will be described briefly in this section.

The base of the accuracy assessment is built by the confusion matrix which denotes the number of true positives, true negatives, false positives and false negatives. Positive and negative are the classes which are change and no change in the case of this thesis. Whether a classification has happened correctly (true or false) is determined in reference to the ground truth on a pixel basis. Consequently, the counts given by the confusion matrix are number of pixels for each case.

Out of this confusion matrix, the following characteristics can be derived:

User's accuracy:	Describes the reliability of the change map, meaning how many of the detected changes are actual changes by setting the correctly classified changes to the total number of classified changes. Also known as ' <i>precision</i> ' or ' <i>positive predictive value</i> '.
------------------	---

$$User's\ accuracy = \frac{TP}{TP + FP}$$

Producer's accuracy: Describes how many of the actual changes are detected by the change map by giving the ratio of correctly classified changes to all actual changes. Also known as 'detection rate', 'recall' or 'true positive rate'.

$$Producer's\ accuracy = \frac{TP}{TP + FN}$$

Related to the two accuracies are also the following two errors:

Error of commission: Describes the ratio of falsely classified changes

$$Error\ of\ commission = \frac{FP}{FP + TP} = 1 - UA$$

Error of omission: Describes how many of actual changes have not been detected.

$$Error\ of\ omission = \frac{FN}{FN + TP} = 1 - PA$$

Finally, the overall accuracy gives a quick overview by building the ratio between all correctly classified pixel and the total number of pixels:

$$Overall\ accuracy = \frac{TP + TN}{Total\ number\ of\ classified\ entities\ (pixels)}$$

Furthermore, another index commonly used in remote sensing to evaluate the performance of a classifier, Cohen's kappa, has been calculated. The kappa coefficient describes the performance of a classifier compared to a random classifier.

Especially Cohen's kappa but also the overall accuracy may have limited explanatory power in general and in specific for this thesis. The kappa coefficient has been discredited increasingly in the scientific community over the past few years. One example of critique to the well-established index would be Foody (2020). He argues about the limited usefulness of the comparison to chance agreement as well as difficulty to interpret the kappa coefficient and the limited ability of comparison. Other authors such as Warrens (2015) and Vieira et al. (2010) attribute at least some usability to the measure developed by Cohen. Due to its prevalence on the one hand

and the controversy on the other, it was decided to include but not solely rely on it for this thesis.

As a further problem, Foody (2020) states that the kappa coefficient is especially challenging if variations in class abundance occur. This is also the biggest limitation when using overall accuracy. In this thesis, the class ‘no change’ vastly outnumbers all change classes. Nevertheless, a universal accuracy measure is needed to allow for a quick-look comparison between different adjustments in the methodology as well as different datasets. This is why further measures to describe the accuracy based on the data from the confusion matrix will be explained below and used for this thesis:

F1-score: Combines User’s and Producer’s accuracy in a harmonic mean (Hripcsak and Rothschild, 2005).

$$F1 - score = 2 \frac{UA*PA}{UA+PA}$$

Also the above-introduced F1-score is not an undisputed perfect measure. Similar to other measures discussed above, it is said not to be totally accurate in cases of imbalanced classes with the danger of overoptimistic results (Chicco and Jurman, 2020, Chicco et al., 2021). Instead, the Matthews correlation coefficient (MCC) is suggested, since it considers all four categories in the confusion matrix and only shows good results if all of them show reasonable values and relations. It is given by the following formula and also known as the ‘*phi coefficient*’:

Matthew’s correlation coefficient:
$$MCC = \frac{TP*TN - FP*FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}}$$

2.4.1. Ground truth

A crucial element of the accuracy assessment is having a reliable ground truth. Real changes only happen between the circular acquisition 0118 and the two linear flights 0110 and 0111 in Memmingen since only one flight has been made in Friedrichshafen. The ground truth has been created by overlaying two 2D SAR images over each other and manually assigning each pixel to the right category.

A depiction of the ground truth can be found in Figure 3.

2.5. Validation with independent data set

For testing and validation purposes as well as to verify if the algorithm has not been overoptimized for the Memmingen example, the Friedrichshafen dataset is used. Because of inferior flying conditions during the acquisition, i.e., wind and turbulence, only part of the data set can be used. This offers limited opportunities to check the performance under different viewing angles, as only a small range of 20 degrees is available for the evaluation.

As a consequence, the maximum available deviation of the angle is used as well as a smaller angle of 8 degrees to check the algorithm's performance under small angles. Since we only have one flight of which all subapertures have to be assumed to be recorded almost simultaneously, we have to work under the supposition that no change has happened. This also implies that all detected change would be false alarms and that no change would theoretically be an optimal result. Due to these limitations, the results of this dataset can be found in appendix D.

3. Results

In this chapter, the performance of the baseline methodology in 2D mode is quantified. The assessment is always in consideration of the idea of varying illumination angles. Results of the Find Pairs algorithm are then introduced and compared to the baseline.

3.1. Accuracy assessment of the baseline method

3.1.1. 0118 vs. 0110

Figure 4 already revealed some of the characteristics of the baseline method. There, it can be observed that with a very small deviation of the angle between the acquisitions, at least visually, the detected changes match the ground truth. By increasing the angle, the number and magnitude of false alarms increase.

The airplanes on the tarmac as well as the presumptive baggage carts are recognized as change in all three images, with the airplanes being much more pronounced in C and less so in B.

As for the buildings and trees, we can see that there are almost no registered changes in A while B features some outlines of the buildings marked as changes. They occur particularly at the edges of the buildings as well as at some trees.

In subfigure C, false alarms are very pronounced with most trees occurring once as a positive and once as a negative change. The same applies for edges of buildings which are marked by one positive and one negative change, usually opposite of each other.

Looking at Figure 12, we can see that the overall accuracy decreases with an increasing angle until the middle of the circle is reached at around 180° and starts to increase again as we return to the origin from the other side of the circle.

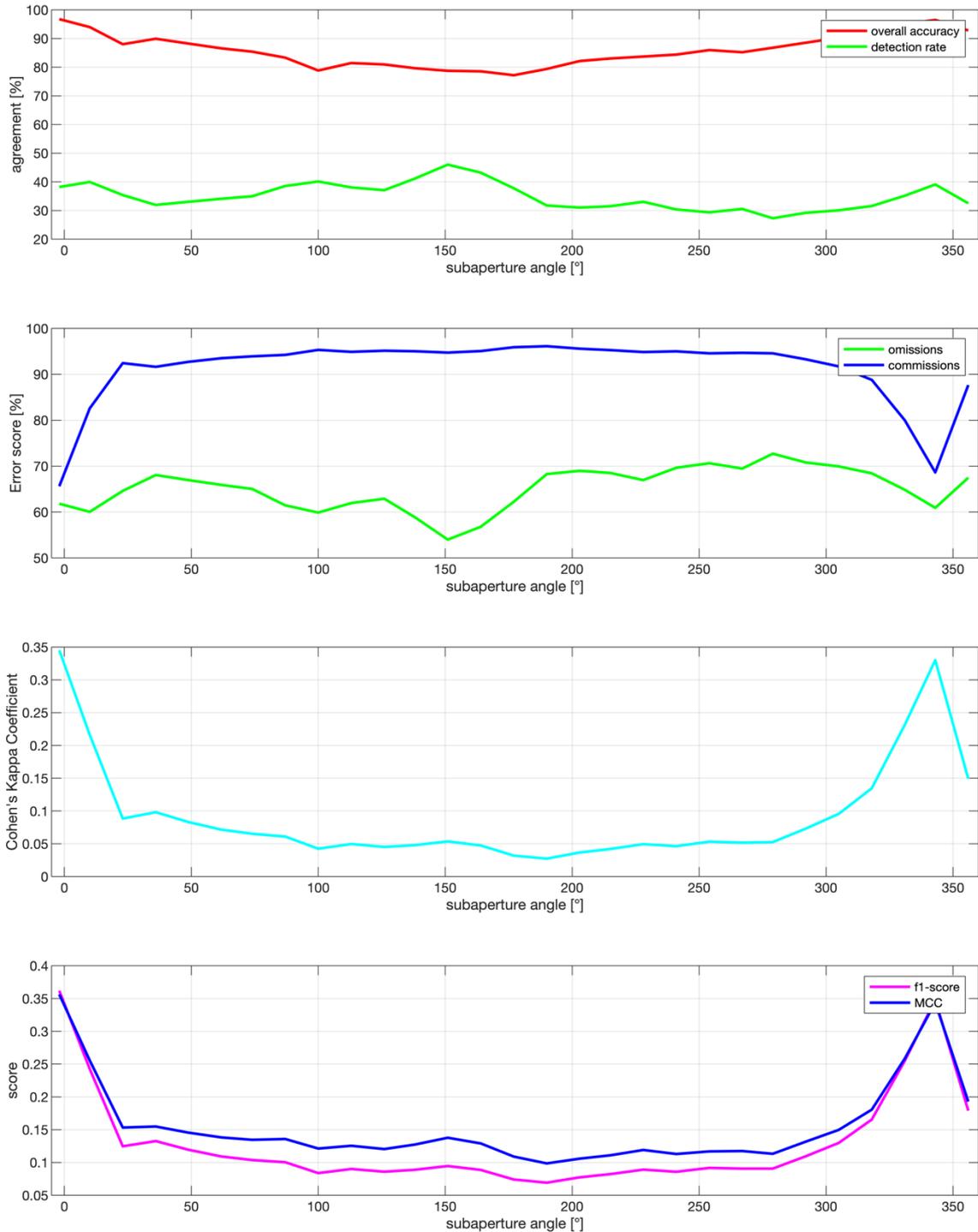


Figure 12: Accuracy assessment of the baseline method showing overall accuracy, detection rate, commissions and omissions as well as the kappa coefficient, F1-score and Matthew's correlation coefficient. Data from circle 0118 against linear path 0110.

The detection rate does not show a clear pattern or trend and fluctuates between 30% and 45%. It reaches its maximum of 46% at around 150° which is not close to the origin, but the closest point to flight 0110. Subapertures closer to the linear path feature a detection rate between 35% and 40%.

The error of commission is consistently high at more than 90% for more than 20 of the 29 subapertures, meaning that it is at or above 90% when having angles greater than $\sim 25^\circ$. The first subaperture features a significantly lower but still high error of commission of roughly 65%. Also the subapertures around 340° have a lower error of commission. Since the error of omission is the inverse of the detection rate, it follows the same pattern and will not be described further.

The kappa coefficient follows a similar pattern as the error of commission, just inversed. It has higher values of around 0.35 in the subapertures close to the linear flight path and is consistently low around or below 0.1 for all other subapertures.

The F1-score follows this pattern as well with the best subapertures reaching values slightly over 0.65. The Matthew's correlation coefficient depicts a very similar result as the F1-score with its highs not as high as in the F1-score and the lows not quite as low. The values range between slightly below 0.4 and around 0.2.

In summary, one can say that most accuracy values either show no clear signal – such as the detection rate and the error of omission – or follow the tendency that they are better below 25° and above 325° , which are close to the linear flight path from which the test image is retrieved. If in this second pattern, the values drop rapidly after the before mentioned subapertures and remain low for all other subapertures. The only measure which shows a more gradual increase and decrease is the overall accuracy.

3.1.2. 0118 vs. 0111

Results of the change detection using the southern linear flight path (0111) can be viewed in the appendix A as Figure 18. The best overall accuracy has been obtained at an angle close to 0° with 95.1% which features the lowest detection rate at the same time.

For almost the entire circle, the error of commission is at or above 95% while only the closest subapertures feature lower values. Also around the 0° angle is the only spike of the otherwise flat curve of the kappa coefficient. There, it reaches a maximum value of 0.11 while fluctuating between just under 0.02 and 0.03 elsewhere. A similar behavior can be observed for both the F1-score as well as the MCC.

In general, it can be concluded that the results of the southern linear track follow a similar pattern considering the behavior of values relative to the angular distance from the subapertures to the linear track. Except for the overall accuracy, the agreement values are however consistently lower and the error scores consistently higher.

3.1. Accuracy assessment of change maps after the application of the Find Pairs algorithm

3.1.1. 0118 vs. 0110

Figure 13 reveals the performance of the Find Pairs algorithm in comparison to the baseline method. The overall accuracy of Find Pairs is consistently above the baseline method. Compared to the initial state depicted in Figure 12, we can see that the overall accuracy assumes consistently very high values of around 95% with a maximum of 97.6% while never dropping below 88%. As such, it offers an improvement in the range of 10 – 15 percentage points over the baseline method. The detection rate fluctuates more and features a similar maximum as in the original approach at the same subaperture but tends to have slightly lower values. The error of commission has a similar pattern as in the original but also fluctuates more.

It could be reduced consistently by using Find Pairs with improvements typically ranging between 10 and 15 percentage points and extremes as high as more than 20 percentage points.

The three measures that try to give a combined statement about all aspects of the accuracy follow a similar value range as in the original albeit having slightly higher values but fluctuate more.

Table 2 exemplifies the findings stated above. It offers a more detailed insight into an example of a subaperture close to the linear flight and one with a greater angle. One can observe that some of the detection rate is sacrificed for a significantly higher overall accuracy and lower commission error. It further shows that at close angles, the result of Find Pairs is not greatly different to the baseline method, whereas this gap grows with larger differences in the angle.

More detailed graphs showing only the measures after the application of Find Pairs can be found in the appendix in Figure 19 and Figure 20. Figure 21 shows the same

analysis as Figure 13 but with flight 0111 as test data source. All observations made in this section are generally also valid for 0111 despite fluctuations tending to be slightly higher.

Table 2: Confusion matrix and accuracy measurements for two different subapertures before and after the application of Find Pairs.

0118-1000 vs. 0110 -2° – Baseline		algorithm	
		change	no change
ground truth	change	6376	10317
	no change	12164	662785
overall accuracy: 96.75%	F1-score: 0.362	detection rate: 38.20%	commission error: 65.61%

0118-1000 vs. 0110 -2° – Find Pairs		algorithm	
		change	no change
ground truth	change	6000	10693
	no change	4570	670379
overall accuracy: 97.79%	F1-score: 0.440	detection rate: 35.94%	commission error: 42.24%

0118-82 000 vs. 0110 113° – Baseline		algorithm	
		change	no change
ground truth	change	6352	10341
	no change	118040	556909
overall accuracy: 81.44%	F1-score: 0.090	detection rate: 38.05%	commission error: 94.89%

0118-82 000 vs. 0110 113° – Find Pairs		algorithm	
		change	no change
ground truth	change	5549	11144
	no change	28124	646825
overall accuracy: 94.32%	F1-score: 0.220	detection rate: 33.24%	commission error: 83.52%

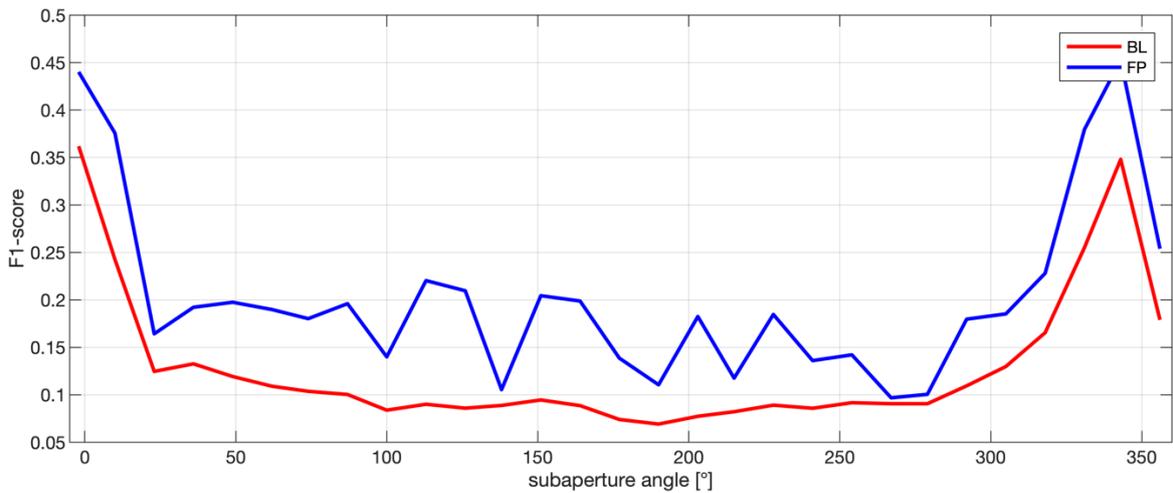
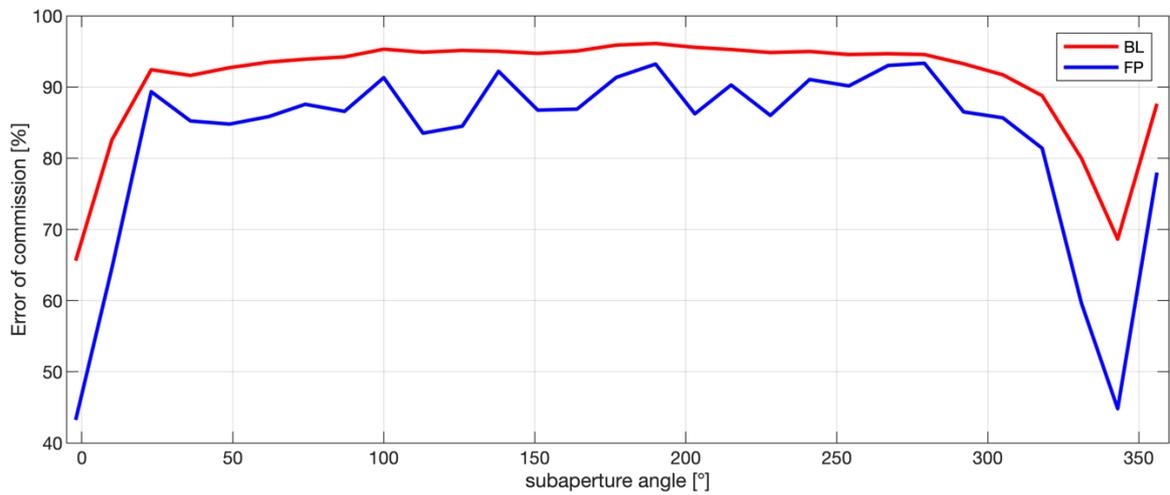
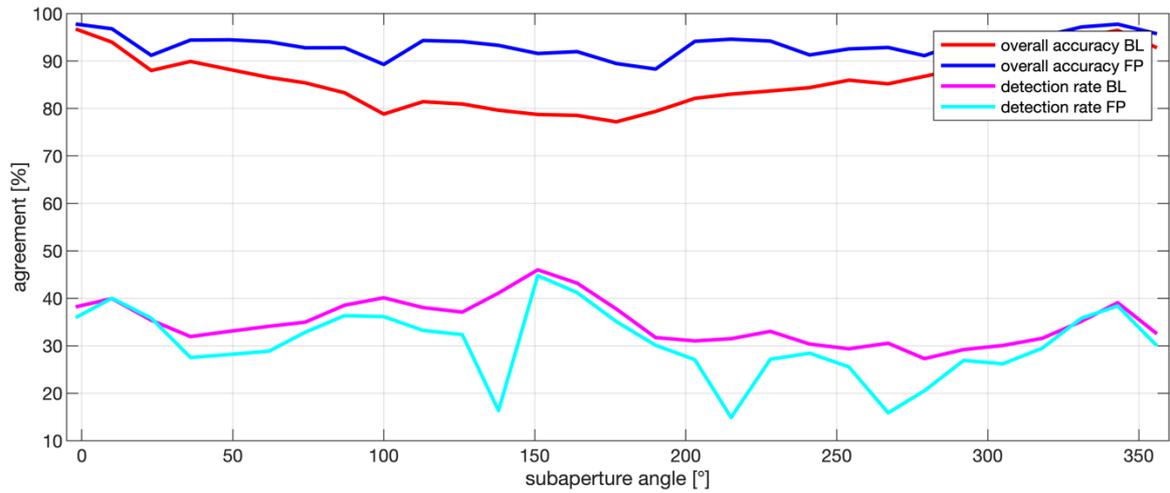


Figure 13: Comparison of the classification performance of the baseline method to the Find Pairs postprocessing. Subapertures of dataset 0118 as reference images and 0110 as test image.

3.2. Visual comparison and performance assessment of both methods

3.2.1. 0118 vs. 0110

Figure 14 offers a visual insight into both methods and their performance in comparison. When looking at the first subaperture, one can observe that there is little difference between the two methods.

Baseline Method	Find Pairs	Subaperture
		#1 first pulse: 1000 -2°
		#3 first pulse: 19000 ~23°
		#10 first pulse: 82000 113°

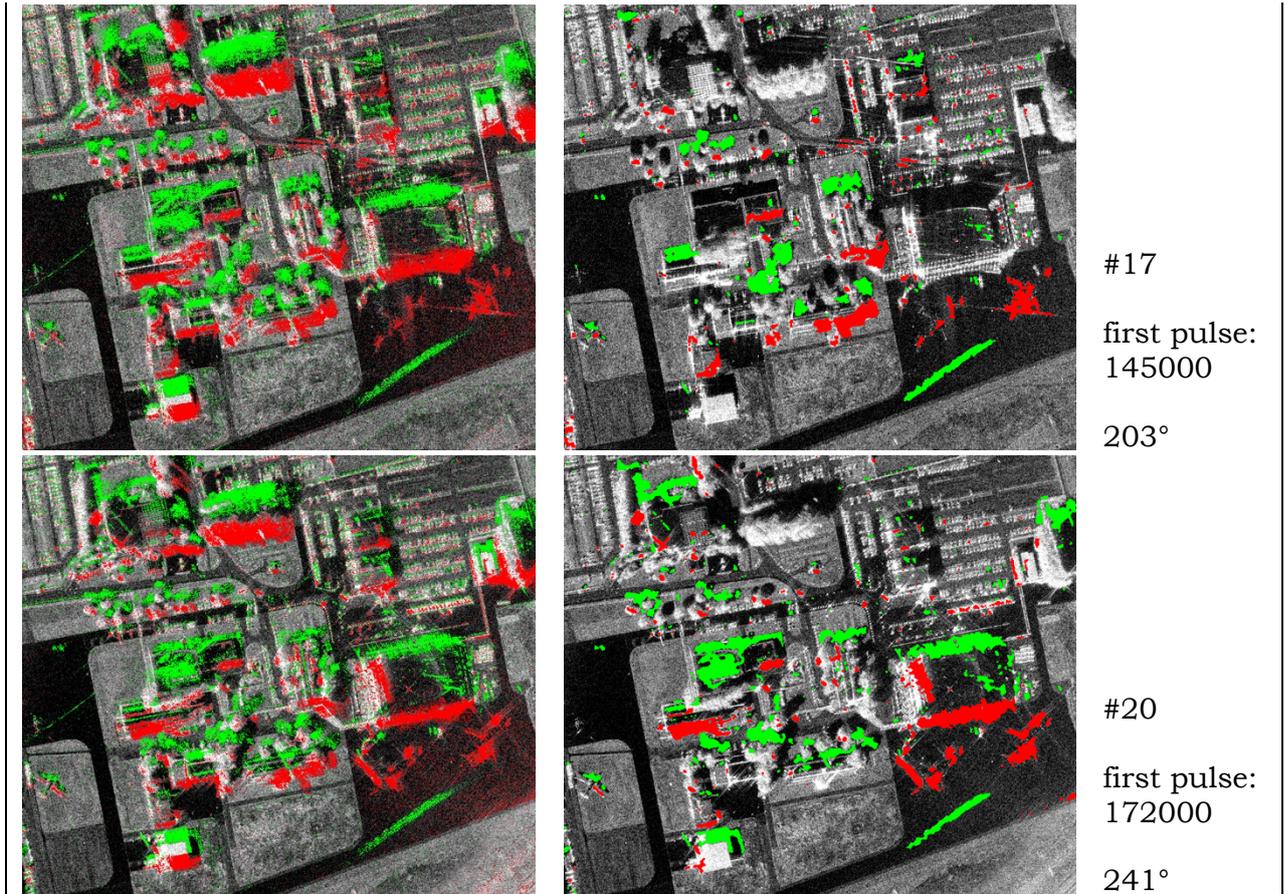


Figure 14: (above & previous page) Comparison of change maps between baseline method and Find Pairs algorithm for data from all subapertures of 0118 evaluated against 0110. The subaperture column denotes the number of the subaperture, their respective first pulse and the angular difference between the reference and test image.

The registered changes coincide in magnitude and their distribution. Many of the registered changes in the parking lot have disappeared by applying the Find Pairs algorithm. Smaller patches of changes distributed throughout the image that only occur in either one of the two change classes can be seen in both variants. Examples are linear and elongated artefacts of positive in the left lower half of the image as well as the crescent-shaped positive change which corresponds to a series of ground service vehicles on the tarmac which is a true positive. Small patches of negative changes mainly persist in the upper half of the image. Other than that, both methods deliver change maps that reflect the ground truth to a great part and are very similar to each other.

With an angular deviation of 26° , the second example already shows pronounced changes detected by the baseline methodology. They cover a significantly larger fraction of the image, and it is apparent that the majority of the changes are of the negative change class. As such, most changes do not have a counterpart of the

positive change class. Find Pairs retains most of the changes from the baseline method although it does aggregate the elements to more defined instances of change and removes noise.

When looking at the third example featuring a 113° difference, we can see that the baseline method forms the typical pairs that have been mentioned before. They occur again at edges of buildings, trees and groups of trees. Most of these changes have been canceled by the new method. Only a pair of L-shaped false alarms caused by a building slightly left of the center of the image still persists. Above these retained changes by the building, a row of single trees forms pairs of changes as well. Most of them have not been removed by Find Pairs although they are not registered in the ground truth. In the top left corner, the baseline method forms a continuous line of a positive change that is partly mirrored by two or three shorter and thinner lines of negative change parallel next to it. After the application of Find Pairs, one of the corresponding negative changes has been removed while the other negative changes are still visible, and the larger positive change has been retained.

The parking lot in the top right corner shows registered changes for most cars with the negative change class being predominant. When compared to the ground truth (ref. to Figure 3), many of these changes are actually false positives. Find Pairs removes some of these changes, but it cannot be said that the removal targets solely the false positives. Change elements appear to be merged compared to the baseline method such that multiple cars are covered by one change element whereas they can still be identified as single elements in the original.

The fourth example shows a similar situation when observing the output of the baseline method. The number, magnitude and pattern of false alarms closely reflects the example before while the angular distance has increased by almost 90° degrees. Applying Find Pairs yields in some different results. Many pairs are removed again while some changes around the center of the image persist. There, two patches of changes of similar size and form and opposed class are noteworthy. The formation of pairs for this building by the baseline method has been different in the previous example resulting in a removal that has not happened at this subaperture.

Again, not all changes induced by trees have been removed, even though, at least visually judging, pairs have been formed in the baseline method.

Looking at the fifth example, it is noteworthy that a hangar right of the image center forms a pair of positive and negative change which was not get removed with Find Pairs. Generally, in this example, there are several pairs that are seemingly related to each other when judged visually but were not removed by the algorithm. This includes again predominantly trees while other remaining change features are mainly single class. An exception would be the lower left corner, where a smaller hangar forms two elements of similar shape and size of which however only one is removed by Find Pairs. Elsewhere, many pairs are again removed analogous to the examples before while all aircraft changes that are also registered in the ground truth have been retained.

In summary, one can conclude that the examples above have shown that applying Find Pairs leads to the removal many registered changes. In case of the examples studied in this thesis, the removed changes have almost exclusively been false alarms when compared to the ground truth. Despite the removal of a significant part of such false alarms, some elements of change that are not found in the ground truth still remain in the mask. The number, position and characteristics of these remaining changes vary from subaperture to subaperture. In some they can be visually linked to other elements that might be subjectively judged to be similar by the human eye while others are more reflecting stand-alone elements with no similar shape of the contrary change class in vicinity.

On a further note, the baseline method confirms the expected behavior of building more false alarms with greater angles, both in number and size. Smaller angular differences have less pronounced objects of false alarms ranging from noise-like features to smaller artefacts with an imbalance between change classes while larger angles bring clearer pairs mostly reflecting the shape of the object causing a pair.

3.2.2. 0118 vs. 0111

Figure 15 shows a selection of change maps for given subapertures and compares the baseline method to Find Pairs. In the first example featuring subaperture 14 with first pulse 118000, which is closest to the southern linear flight path 0111, we can see that the baseline method has detected the two airplanes as well as some artefacts right to the center of the image.

Example 2 features an angular difference of 33° , meaning it is only 2 subapertures apart from the closest. Nevertheless, it already shows many mirrored changes that are not registered in the ground truth. They consist of narrow patches at the edges of buildings and other structures. Find Pairs removes few of these features but changes their appearance such that they emerge more crumbled than in the original. The airplanes are registered in both versions, and it is noteworthy that at least one of the two airplanes is listed both as negative and positive change. Their labeling does not consist of a contiguous change element but are composed of multiple smaller fragments that are aligned along the fuselage and the wings.

The third example (subaperture #17) shows very prominent elements of change most of which have a mirrored counterpart of the opponent change class nearby. Many but not all of them are removed by the find pairs algorithm. Also some changes in the parking lot disappear under its application. Interestingly, one of the two airplanes is now registered mostly as positive change, while it has been mixed and only negative in the examples above.

A rather large element of change around the large hangar above the planes which actually has a counterpart on the other side has not been entirely removed by Find Pairs.

The last example has a similar deviation in the illumination angle but in the other direction. Like in the above explained case, it shows many elements of change not to be found in the ground truth that are mirrored in many cases. However, the negative change class predominates the scene slightly both in frequency and magnitude. Find Pairs managed to remove many of those changes.

When regarding the apron, however, we can see that the baseline indicated extremely thin parallel lines of change speckled with some noise and single pixels of change. Find Pairs appears to have amplified those lines by again creating crumbles that are much more pronounced than the lines found in the original. These artefacts overlay the airplanes and potential changes that could have been indicated there such that it remains unclear if these changes have been detected in this example with Find Pairs. Also in the baseline method, the planes are not marked very clearly but are at least hinted by the change detection.

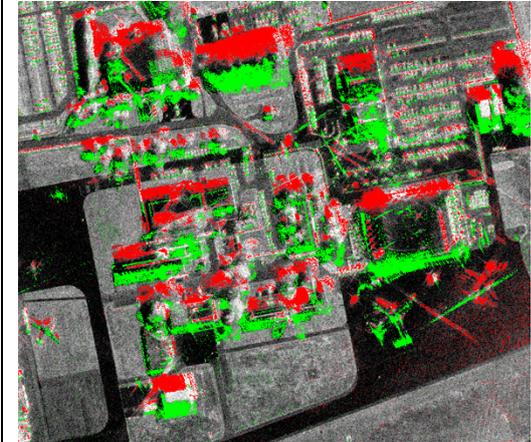
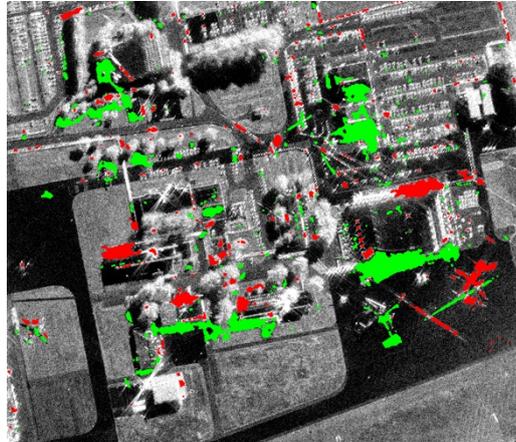
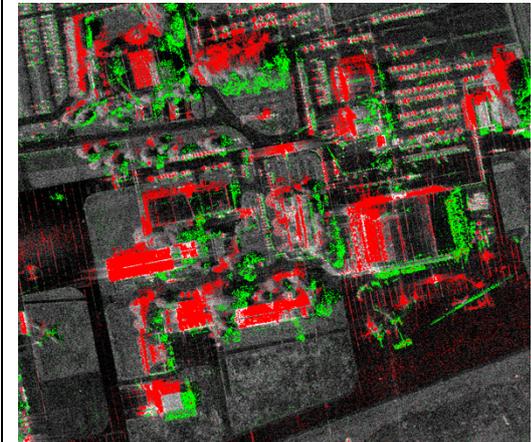
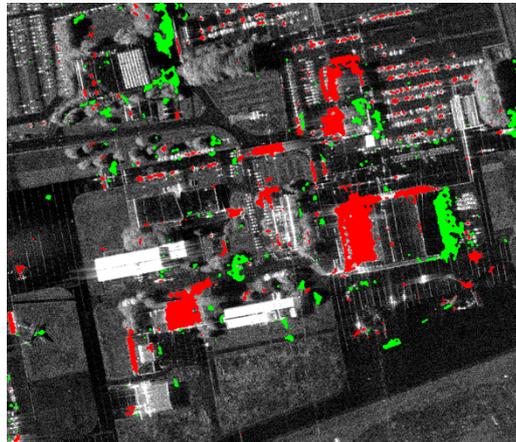
Baseline Method	Find Pairs	Subaperture
		#14 first pulse: 118000 -7°
		#12 first pulse: 100000 -33°
		#17 first pulse: 46000 -110°
		#23 first pulse: 199000 108 °

Figure 15: (previous page) Comparison of change maps between baseline method and Find Pairs algorithm for data from all subapertures of 0118 evaluated against 0111. The subaperture column denotes the number of the subaperture, their respective first pulse and the angular difference between the reference and test image.

In summary, one can say that Find Pairs managed to remove many changes not consistent with the ground truth also in this dataset. Some phenomena, like the building of crumbles, seem to occur more frequently in this dataset. While it did improve the change detection at least to some extent in most examples, the results seemed not to be as consistent. Many artefacts did get concatenated or in a few cases even slightly amplified through Find Pairs. On the other hand, the change maps retrieved through the baseline method also varied in their behavior of building pairs and showed differences in their characteristics such as the thin lines as well as alternating change classes for the airplanes.

3.3. Weighting and sensitivity to the different agreement criteria

Observing the agreement values of different criteria for selected pairs has shown different patterns. The empirically judged best result for the algorithm could be reached with the weights listed in Table 3. Higher weights have been used for area and distance, since they appeared to give a reliable indication about the probability that two pairs belong together. Height, standard deviation as well as the average gray value showed a mixed signal to the finding of pairs such that smaller weights have been chosen for these criteria.

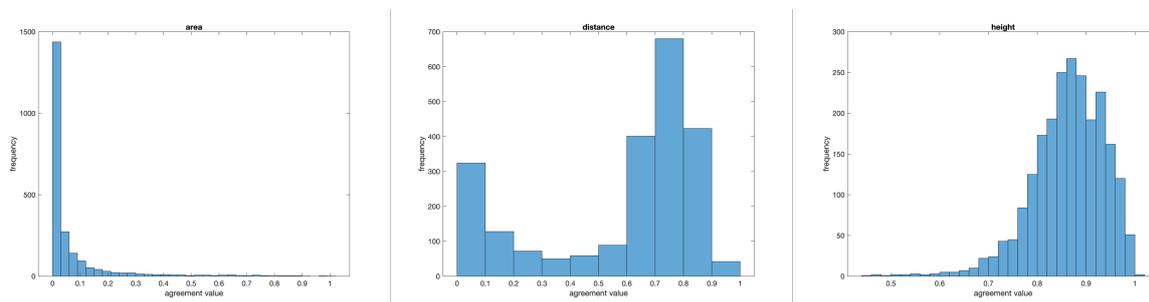


Figure 16: Distribution of different agreement values for pairs in change detection between 0118-145000 and 0110.

The different criteria do not only show different impacts on the selection of pairs by the algorithm, but they also have a big variety of the distribution of values, as can be observed in Figure 16. The agreement in the area is heavily skewed toward lower

values, while the agreement in height only shows relatively high values. The agreement in distance, shows peaks in frequency both at low as well as high values. Agreement values for the geometrical criteria circumference, compactness and number of changes in direction have a skew comparable to the agreement in area.

Table 3: Weights used to retrieve total agreement

agreement criterion	weight
area	3
distance	2
height	0.5
circumference	1
compactness	1
No. of changes in direction	1
standard deviation (contrast)	0.25
average gray value	0.5

4. Discussion

This chapter will address the findings of the previous chapter and comment on the implications that can be drawn of the results and tries to show possible explanations for the behavior of the algorithms used and developed. Furthermore, limitations in the chosen methods are described and their functioning is critically reviewed under consideration of possible scenarios of application. Finally, opportunities shall be highlighted and potential for further enhancements will be outlined.

4.1. Evaluation of the baseline method's results

The results in Section 3.1 show what has been indicated several times throughout this thesis: the change detection is generally dependent on how similar the acquisition geometries are. If this is not given, problems arise and the reliability of the change detections degrades rapidly with an increasing difference in the angle. In Section 3.1 we learned that while the overall accuracy is at a reasonable to high level with roughly 80% and more, other indications show the shortfalls of the results.

A detection rate hovering around the 30 – 40% would imply that more than half and up to two thirds of all changes remain unrecognized. Here, it is important to notice, however, that the accuracy assessment as conducted in this thesis is a pixel-based approach. Looking solely on the actual objects of change, i.e., the airplanes, one could conclude that all objects of change have been registered, albeit not in full extent or maybe not matching the exact footprint as delineated in the ground truth. Nevertheless, this would imply a theoretical detection rate of 100%.

Considering the results both visually as well as statistically, one might deduce that the flaws of change detection under increased angular distance is more a problem of committed classifications rather than omissions. This idea is amplified by the fact that in most subapertures, well over 90% of the indicated changes are false alarms. Again, the application of a strict pixel-based evaluation is at least worth discussing. Omissions on the other hand are a secondary problem in comparison, despite the numbers not being excessively high due to the pixel vs. object problem. Looking at the main change features, i.e., the airplanes, one can see that they have been reliably detected in all subapertures, at least for flight 0118 vs. 0110. This can be explained by the fact that the baseline method mostly detects all real changes at a near to zero

angle difference and detected changes tend not to disappear with increasing angles. They may be overlapped or displaced by effects of false detection to some extent, but this effect seems to be statistically irrelevant. As a result, the detection rate and error of omission do not degrade with an increasing angle, as opposed to most other accuracy measures. It is remarkable, however, that already angles greater than 25° show an error of commission close to the maximum. Judging visually, the fraction of the image area covered by false alarms is not already as high at these lower angles.

The detection of changes of the cars seems to be more challenging. When comparing the results of different subapertures to the ground truth, one can recognize that the change maps are not consistent regarding their statement about the cars. Furthermore, while the subapertures of small angular difference may deliver a result that marks roughly the same cars in the same change class as registered in the ground truth, starting from small deviations, the detected changes by the baseline method seem not to follow any pattern or rule. Instead, they seem to be random. Even in the best case, the cars are only marked partially if at all, evoking again the pixel vs. objects problem. In larger angles, the cars are marked more clearly but inaccurately.

The problems described in the above paragraph are attributable to the size, height and shape of cars. When observing the 2D SAR images without change mask (e.g. Figure 3), one can see that cars are recognizable but one needs to assume that they are only perceptible as such due to background knowledge and the association with the parking lot and roadways. Their limited size of a few countable pixels in the example of this thesis would probably make them unidentifiable as a standalone feature. Despite the brighter backscatter compared to the paved background, their signal could just as well be noise. This becomes more apparent when looking at the parking lot in the top left corner, where single cars are clearly hard to be identified. Given the distribution of the pixel values in such areas and considering the methodology using expectation maximization and thresholding, it is understandable that the performance of the change detection algorithm is limited for objects of this characteristic.

The appearance of false alarms in form of pairs as it is addressed by the algorithm developed in this thesis is most prominent with larger angles. False alarms are however not limited to such occurrences in pairs. Especially in smaller angular deviations, other patterns emerge as well while pairs are less prevalent. Smaller

angles also lead to situations where either the positive or the negative change class outnumbered the other while in large angles, the two change classes are somewhat balanced, also facilitating the good performance of Find Pairs in greater angles.

Looking at the accuracy measures, it is noteworthy that the detection rate is best at one of the largest angular deviations where all other indications of accuracy are low. This can be explained by the fact that greater angles tend not to reduce the detections found at smaller angles but may cover more of the actual changes by overlapping false detections to real changes. As such, this higher detection rate cannot be awarded to the detection itself but rather on a chance agreement.

Further worth mentioning is the high statistical difference in the accuracy measures between the circle and the two linear tracks 0110 and 0111. Possible explanations include differences in flying conditions between the two flights, which may be surprising given the small spatial and temporal gap of only five minutes between the recordings. Data from the inertial measurement unit do not reveal any significant difference in the aircraft's attitude during the recording. Another explanation is the fact that the ground truth is based on the changes between 0118 and 0110. However, when judging visually, one might perceive a tendency that the results from 0110 indeed show more accurate change maps in the lower angles.

Another anomaly that can be observed with the southern linear flight is the inconsistency of the change class of the airplanes on the tarmac. This can be partially explained by the fact that the southern linear path had been recorded before the northern track and one of the two airplanes has still been in motion at this time.

4.2. Evaluation of Find Pairs' results

4.2.1. Changes retained by Find Pairs

The example of a subaperture with a small angular difference exemplifies that Find Pairs does not remove change objects if there are no pairs. The results have shown that for very small angles where also the accuracy measures are high, the baseline method already delivers a product that closely reflects the ground truth, at least object wise. Smaller artefacts of false detection are mostly retained as they lack a

close counterpart of the opposing change class. This also applies for larger change elements, since they do not feature any mirrored element as they are true changes.

Another case where Find Pairs does not remove most of the changes found by the baseline method is found at 23° in Figure 14. Contrary to the case of smaller or no angular difference, most changes detected here appear to be false alarms. The reason for this is that no or only very few pairs are built by the baseline method. This raises the question as to why such large and numerous false alarms are built already under a relatively small angle. Although this cannot be explained without uncertainty, a likely explanation is a strong backscatter due to a perfect reflection only present at this angle. This is also hinted in the change maps by very bright spots in the background. The fact that this phenomenon immediately disappears at slightly larger angles speaks in favor of this theory.

The absence of the opposed change class can be explained by the fact that the angle is still too small to form the shadow effects that usually lead to the pairs through the shadow being at a different position in the two images. Due to the smaller angle, the distance in the positions of two potential shadows is not large enough to form the pairs.

Although this example did not deliver the desired result both in the baseline as well as in the Find Pairs method, it is still true to say that the algorithm performed as intended by design. Having elements of change not mirrored by the opposed change class is a strong indication for a true change when assuming that the baseline method is able to retrieve changes correctly under ideal conditions. Assuming that a comparably large change would have happened due to the destruction of a building or the removal of landmass, the representation using the baseline method and a near-to-zero angular deviation would lead to a result that appears analogous to the 26-degree-example. In this case, the removal of any change by Find Pairs would lead to a wrong result.

Other than the special case discussed above, the potential of Find Pairs could be largely exploited. If pairs of similar size and shape within a reasonable distance were prevalent, they have been removed. There are however numerous cases, where pairs have not been removed albeit their seemingly close relation. Examples are trees and buildings as described in Section 3.2, where the knowledge of the object implies an obvious correlation. For some of these examples, the explanation lies in the closeness

of a pair to another pair. Due to the morphological preprocessing introduced in Section 2.3.1, change elements that are very close may be connected to one element. This usually only happens for one of the two elements of a pair what results in a situation, where the merged element is significantly larger compared to its counterpart and also does no longer feature the same shape. This is the case for the trees mentioned where nearby trees form change elements to be potentially merged as well as buildings, where another section of the building or another adjacent structure leads to the presence of a larger or merged change element in one class. A linear feature might become an L-shaped feature, not only increasing the complexity of the structure as such but also changing the circumference and compactness and any other geometrical measure. This inevitably yields in a lower agreement value thus inhibiting a potential removal. Furthermore, assuming otherwise perfect pairs in the local surrounding of such a phenomenon, the number of elements of each class would not match anymore. Consequently, one element needs to be omitted from the removal process.

The change detection between with the southern linear track revealed another possibility how a pair may be at least partially omitted from the exclusion. Some pairs are not completely connected either by the baseline methodology or by a division that could happen as a result of the morphological preprocessing. When extracting the objects, one element of a pair might be subdivided in two separate objects. Two of the three resulting elements may or may not form a pair; either way, at least one element of false detected change remains. If the difference in area gets too large, it can happen that all elements are retained. Also the geometrical properties are likely to be altered through such a division and can hinder a removal if the impact on the agreement value is large enough.

4.2.2. Agreement in height as an example for the impact of the threshold

Besides the cases discussed above, there are occurrences of similar elements in close distance that are mirrored and not matched with adjoining patches. To investigate on such incidences, the process will be elaborated further for the example of the subaperture featuring an angular difference of 236° . The change map of an area with a pair that has not been removed by the application of Find Pairs is displayed again in Figure 17 (A).

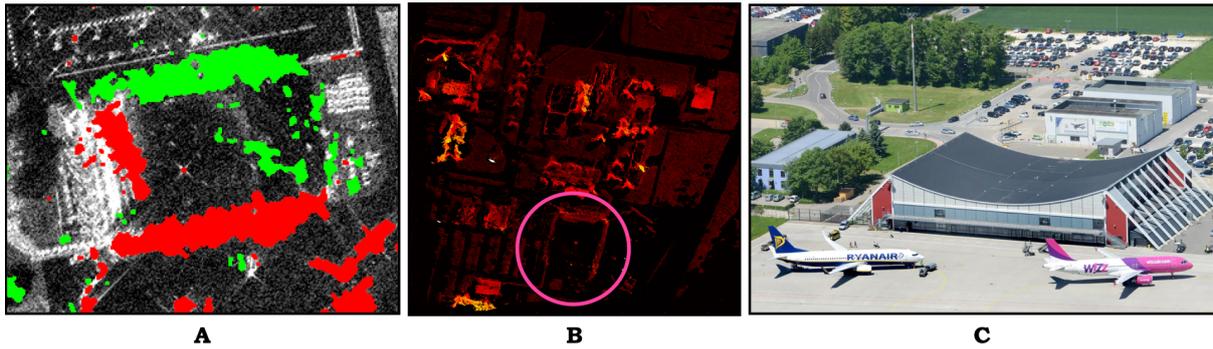


Figure 17: Example of a building creating a pair that is not removed by the algorithm in 0118-172000 vs. 0110 (236° difference). Change map after the application of Find Pairs (A), height of objects determined by reference and test data (B), RGB image of corresponding building (C, source: Augsburg Allgemeine)

A check of the intermediate results of the algorithm shows that the two pairs have been assigned to each other during the finding of a best match. The agreement matrices show very high agreement values for all geometrical criteria as well as the standard deviation. In total, the pair reaches an agreement value of 0.85 which is slightly above the set threshold of 0.8, which is why the pair was not excluded from the change map. One could deduce that an increase in the threshold would easily solve this shortcoming. However, this would also increase the risk of committing changes as false positives that are real changes. It has been decided to leave the threshold at this conservative value, since cases like this occur in a relatively low frequency. The question remains, however, how the building in Figure 17 is creating a pair in many subapertures that usually features such a high agreement that both elements get excluded from the change map without showing the same behavior in this subaperture. Visually judging, both elements have obvious similarity what is confirmed by most agreement values. The reason actually lies in the agreement in height. The building features a special structure because of a half-pipe shaped roof (ref. Figure 17 C). The SAR reflection of this building varies greatly with changing angles, which is why the height information cannot be retrieved as reliably as for more regularly shaped structures. This is exemplified in Figure 17 C, where one can observe low height values for the building while other structures such as the buildings in the background as well as many trees feature much higher height values despite not being taller. Also the point cloud in Figure 1 derived from all subapertures of the circle shows that the height profile of this building cannot be reconstructed nearly as good as any other tall object in the scene.

Since the agreement in height has proved to show a mixed reliability, a lower weight has been chosen for the agreement.

4.2.3. Removal of changes without the presence of pairs

On the contrary, Find Pairs showed to remove elements of change in some cases where visually, no obvious pair could be found. Examples would be change maps where either of the two change classes is dominant and Find Pairs still removes a great part of the false alarms. This includes the last subparture in Figure 15 where not all removed elements have a counterpart that one might assign based on a visual judgement. Here, a deeper analysis revealed in two examples that Find Pairs matched two sets of elements that feature a relatively large difference in their area but have an otherwise great agreement in the other criteria. Despite the higher weight used for the area, the total agreement value still is below the threshold such that the change elements get removed. This is right in this case but should not happen since they do not really match, which could be a true change in another example. A likely reason for this behavior is that through the normalization, the bad value of the unmatching area gets weakened by other pairs that have values that are even worse. Due to the predominance of one change class over the other, which is not only expressed in a smaller number of elements but also by a generally smaller size of objects, there are significantly fewer pairs of comparable size. As a result, pairs that have a good agreement in other criteria are matched despite a mediocre agreement in the area, decreasing the expressiveness of this otherwise important criterion in such cases.

There is another situation where Find Pairs conducts significant removal of objects and pixels of changes without the presence of pronounced pairs. Especially when looking at the results from the Friedrichshafen dataset (ref. Appendix D), one can see that not all changes between the baseline method and Find Pairs are attributable to the finding, assigning and exclusion of pairs of suspected false alarms. In this case, it can be explained by the morphological preprocessing steps conducted in preparation to the actual pair finding. Multiple patches of low density may be concatenated to single, larger objects which are then more likely to be excluded in the pairing process. Furthermore, the preprocessing also removes pixels of change entirely, which brings a denoising effect that may be large in examples of small, dispersed change elements. On top of that, the effect of the predominance of many bad pairs leading to a certain “upvaluation” as described in the previous paragraph may come to play as well. Depending on the compilation of elements in a change mask, the effect of this smoothing and denoising functionality may be more or less pronounced. One must keep in mind that this can also lead to the exclusion of true

changes. Small objects, such as cars in the example the Memmingen dataset, are especially susceptible for this.

On the other hand, this smoothing and denoising effect is the likely explanation for statistical improvements for examples, with similar angles which already exhibit good results. Here, a few pixels of artefacts and noise are reduced through the preprocessing without impacting the detection of true changes, thus building a desirable effect.

Another effect that is most likely to be attributed to the preprocessing routine is the crumbling effect that could mainly be observed in 0118 vs. 0111 and also in the Friedrichshafen data set. Thin change elements assume a more compact shape which has a mixed impact on the quality of the resulting change map. On the one hand, it can facilitate the removal of false detection by enabling the linkage to other elements but on the other hand, it can also lead to situations where change elements do not get removed and do not make sense anymore even in the context of a false detection since the shape is in no correlation to the causing object.

Generally, the application of Find Pairs leads to a slight reduction of the detection rate. Through an optimal parameter configuration, this reduction could be kept at a low level such that it is outweighed by the gains of the method.

The decrease in the detection rate is caused by the removal of true changes, which mainly happen in the parking lot as well as parts of the airplane and ground service equipment by reducing the footprint of some patches of change, both through the pair removal routine as well as the preprocessing steps.

4.3. Methodological limitations

One of the main difficulties of evaluating classifications such as change detection is the case of unevenly distributed classes (Yi-Min and Shu-Xin, 2005). Also in the example of this thesis, the prevalence of any change class is outnumbered greatly by the class 'no change'. This is exemplified by the fact that the application of a trivial detector leads to an overall accuracy of 97.5%. It must be noted, however, that all other accuracy measures equal zero when using a trivial detector, with the error of omission being the exception with 100%.

This induces several challenges in the usage of different accuracy measures. As already shortly discussed in 2.4, the kappa coefficient as well as the F1-score are prone to misinterpretations and false conclusions especially when dealing with unbalanced classes. It has been shown that different literature reflects differently on these problems but especially the overall accuracy is highly disputed by uneven class sizes. In this thesis, we were able to consistently increase the overall accuracy. Especially in the subapertures with the lower overall accuracy values of around 80%, an increase to more than 90% could be achieved. Although this is a significant positive improvement, it is important to consider the limitations of the accuracy measure discussed above.

All other accuracy measures are consistently improved, although their signal may be subject to more fluctuation. Only the error of omission increased slightly which however has been expected. All this implies that Find Pairs gives a significant improvement over the baseline method that is mainly pronounced in larger angles. Not all accuracy measures confirm the magnitude and consistency of this improvement.

4.3.1. Availability of changes for performance evaluation and ground truth

In this project, only one data set offered a time difference with changes between acquisitions. Even there, the number of changes that have happened is limited to the two airplanes and a number of cars in the parking lot. The latter are very small objects that are represented only in a few pixels in size and are all very similar in shape and structure. Furthermore, the time difference of less than an hour does not allow for any greater changes to happen. Thus, the diversity of these changes is limited to mobile objects such as vehicles and other transportation devices.

As a result, it was neither possible to test the performance of the algorithms on permanent changes such as the construction or destruction of buildings and other infrastructural elements nor on changes of natural features. The latter includes important applications of geographical problems such as changes in vegetation but also structural changes such as hang slides, avalanches and other erosional effects or displacements induced by tectonical activity.

Since only one acquisition was available for the validation, the algorithm could not be tested on its performance on real changes in an independent dataset.

On another important note, this assumption of no changes within the short duration needed to conduct a circular flight must be kept in mind. It can be expected that there are some minor changes between the different subapertures of a circle and also in the five-minute time difference between the two linear acquisitions in Memmingen. Some smearing visible on some images and changes masks may be attributed to objects that have moved during the recording. Therefore, strictly speaking, the ground truth is not valid for all subapertures. For this thesis, the changes within a recording have been assumed to be neglectable and are expected not to lead to a significant alternation in the result. Nevertheless, this simplification needs to be considered.

Speaking of the ground truth, it is also important to declare that this ground has been manually created by the author. As such, it is susceptible to human error and bias since the author has been in contact with the output of the change detection before creating the ground truth.

4.4. Implications on the applicability of the method developed in this thesis

Find Pairs has showed to improve the change detection by reducing false alarms both visually and statistically. The algorithm shows good performance when having pronounced elements of false alarms in large angular deviations. In small angles, Find Pairs might not be needed due to the reduced occurrence of pairwise false alarms. However, it has been shown that the algorithm recognizes the absence of such pairs and does not overcorrect by removing real changes. Implausible removals only occur in special situations which are usually found in intermediate angles where one change class is more dominant, and normalization leads to overestimation of the agreement due to a generally low agreement for most potential pairs. Furthermore, care should be taken when applying the algorithm on noisy change maps that lack any pronounced features of false alarms and have dispersed patches of low pixel density.

The right setting of the parameters and threshold is crucial and also the tuning of the weights is subject to be changed for each dataset. This may increase the workload needed to apply the algorithm, making it especially worthwhile when having a set of multiple recordings for a scene with similar properties, e.g., a circular acquisition,

and maybe less so for single observations. Nevertheless, even in the latter case, Find Pairs can be a viable method to systematically reduce false alarms.

Of course, the improvements achieved come at the cost of a slightly reduced detection rate. In most cases observed, the gain in accuracy through the reduction of false alarms outnumbers this decrease by far.

Find Pairs always requires the previous execution of a baseline 2D change detection method. As such, it can be viewed as an add-on or postprocessing function and has been developed and tested to work best with the method developed by Mendez Dominguez et al. (2018b).

4.5. Outlook

4.5.1. Possible improvements and enhancements to the method developed in this thesis

As a first potential improvement to the method developed in this thesis, the tuning of the parameters needs to be mentioned. So far, a relatively robust, manual approach has been applied using empirical insights gained on only a selection of known pairs, also considering only a certain subset of subapertures. It can be assumed that the weights and parameters used could lead to even better results if more examples and more subapertures were considered. This combined with the great number of different parameters that can be adjusted in the method developed may lead to a high workload. Furthermore, manual parameter tuning also brings the flaw of being susceptible to human error and bias. Consequently, as an alternative approach, an automatized parameter optimization could be conducted provided that a ground truth is present. An algorithm would find an optimal solution by iteratively testing different combinations of values such that the deviation from the ground truth is minimized. An example of an algorithm that could be used would be the Monte Carlo simulation (Kroese and Rubinstein, 2012).

Any algorithm would however be adjusted to the specific problem of this thesis and should allow to define ranges for the parameters to be tuned. Another solution that eases some of the challenges with manual parameter definition but does not include full automatization and would not involve all parameters is the automatic determination of some parameters via image properties. Examples here would be the

distance function as well as the size of the structure elements for the morphological operations, both of which should be dependent on the image size as well as the size of the objects.

Although the combined 3D change detection method developed by Mendez Dominguez et al. (2019) has not been evaluated extensively on its behavior concerning false alarms in this thesis, 3D SAR change detection may be subject to the building of pairs as well. As such, it would make sense to create a version of Find Pairs that can be applied to 3D change detection to remove potential pairs of false alarms there as well.

The results have shown that there are cases where Find Pairs may remove elements from a change map even without the presence of pairs which would not necessarily be assessed as matching pairs. This is mainly attributable to an effect which occurs when a scene predominantly consists of pairs with a lower agreement. There, the normalization leads to an upvaluation of the best of these generally bad matching pairs. Here, it would be beneficial to find a way to avoid this upvaluation. A possibility could be to also involve unnormalized values for the thresholding and use the normalized values for the ranking and assignment of the pairs. Solving this deficiency would improve the algorithm's performance in intermediate angles.

Generally, Find Pairs would benefit from additional testing on other datasets, especially datasets with more or more diverse real changes. Insights gained could be used to further adjust the algorithm's properties and to obtain more certainty about its general behavior.

For this thesis, the threshold for the agreement deciding which pairs will be removed and which will not, has been determined empirically based both on the values of known pairs as well as the distribution of the agreement values by looking at the histogram. Alternatively, the threshold could be retrieved automatically through the frequency of the value by fitting a curve and using the first or second local minimum, for example.

Although Find Pairs has proved generally not to exclude elements of change that are real changes in the dataset used for this thesis, it can be made even more resilient to a potential removal of a true change. This could be achieved by including the information of the known angular deviation between the reference and test acquisition. As such, the position of a potential false alarm can be predicted. This would give more information about the likelihood that two elements build a pair of a false alarm.

4.5.2. Alternative approaches

Due to the behavior of existing change detection methodology under increasing viewing angles, the approach using pairs has been chosen for this thesis. Rather than being a standalone change detection method, it postprocesses the output of an existing methodology and has the role of an add-on function. Instead, other methods of change detection could be implemented that are inherently less prone to varying viewing angles.

Such methods could make use of semantical information, e.g. by classifying objects as described by Hugues et al. (2018). Here, 3D point clouds acquired by laser scanning technologies such as LIDAR have been used, but it could be adapted to TomoSAR as well. Change detection then would assess the change in class at a given position.

Semantic classifications also require aggregating the data into objects. As a more general concept, it would also be possible to do an object-based change detection rather than a pixel/voxel-based approach. Similar to some of the methodology used in this thesis, size, texture and shape information is used to segmentate an image to create objects. This contextualized information is then used to detect changes instead of conducting a classification (Ban and Yousif, 2016). Besides reducing computational complexity, object based change detection is reported to be less sensitive to errors in image coregistration (Ban and Yousif, 2016). Despite the fact that we assume correctly coregistered images for this thesis, the effects caused by different viewing angles may be somewhat comparable to coregistration errors. Coregistration encompasses the process of spatially linking pixels of different images to make sure that the same area on the ground is registered in pixels that are considered to map the same area (Li and Bethel, 2008).

On a further note, Guo et al. (2021) have applied a density-based clustering method to retrieve buildings from TomoSAR point clouds. They also mention the difficulty of noise and false targets as one of their motivations. Their method could be developed further such that it is able to generally extract any shape or structure, such as vehicles, trees and other natural objects and be used for change detection purposes.

4.5.3. Potential fields of application

Besides the rapid mapping and emergency response purposes already mentioned in the introduction as well as urban and infrastructure studies, geography offers a great

variety of more potential applications for the method developed in this thesis. The algorithm has proved to show a solid performance on the removal of false detection caused by trees. As such, it could be used for forest mappings and changes in the forest inventory, similar to the applications described by Yu et al. (2015) and (Durieux et al., 2019). In recent decades, challenges induced by climate change such as droughts and wildfires, but also invasive species have impacted forests on a variety of magnitudes and time scales. Data availability is crucial to gain insight on the characteristics of such problems such that they can be addressed appropriately. Here, Find Pairs can help to exploit data from recordings that have not been initially coordinated for their interoperability.

Similar to the data used for this thesis, other studies using the method could reveal changes in land use, especially if changes between urban and natural classes take place. In the wake of the increasing frequency of heat waves, many cities endeavor to find strategies to reduce the suffering of their local population (Huttner et al., 2009). Here, Find Pairs could help to quantify the effect of planting or removing trees along infrastructural elements on temperatures on a micro scale. A further example would be changes in the Swiss landscape. Agricultural decline combined with urban sprawl have led to a continuous increase in both urban and forest area on the expense of agricultural land (Price et al., 2015). Here, find pairs could help to better describe the process of scrubs displacing alpine meadows followed by the growing of a forest.

5. Conclusion

In this master's thesis, an algorithm to address false alarm in 2D SAR change detection has been developed. It is based on existing 2D change detection methods developed by Mendez Dominguez et al. (2018a) using a difference image, expectation maximization and thresholding to generate a change map and removes false detections occurring as pairs. This is mainly beneficial if a certain difference in the illumination angle between reference and test image creates false positives that are mirrored in both change classes.

The algorithm has proved to show a good performance especially in large angular deviations while it does not alter the result significantly if applied on a set with a small or no angle between the recordings. Overall accuracy could be increased by a range of 10 –15 percentage points bringing greater angles close to the results of the closest angles. Detection rate has slightly decreased in the range of 5 – 10 percentage points for most examples. Challenges persist in special situations usually found in intermediate angles. Nevertheless, the algorithm managed to almost consistently improve most accuracy measures, bringing results from large angular differences much closer to the results gained by applying the baseline method on a set of reference and test image that are recorded from the same viewing angle. As such, the false alarm rate could be reduced significantly while preserving most of the detection rate of the outgoing methodology.

A second set of data did not feature any real changes, such that only a limited validation could be conducted. However, it has shown promising behavior on the false alarms retrieved by the baseline method. Therefore, it would be advantageous to apply the developed method on further datasets to gain further insight and adjust the algorithm's parameters if needed.

The method developed can be widely applied wherever SAR change detection may be beneficial.

Acknowledgments

Without the generous support of several people, this thesis would not have been possible. First and foremost, I want to express my deepest gratitude to both of my supervisors Dr. Daniel Henke and Dr. Elias Mendez-Dominguez. They always offered their time to discuss about any of my concerns and ideas. Daniel helped me to channel my thoughts to build a meaningful project and gave treasured inputs pointing me into the right direction. Elias never rejected any data processing request, no matter how extensive this was and always helped me to better understand both fundamentals as well as more complex concepts underlying the thesis.

Furthermore, I want to thank the entire group of SARLab for giving me such an exciting insight to their work during countless meetings and inspiring coffee breaks and especially Emiliano for offering his support of proofreading during the last days before submission. Here, I also want to mention my sister Christina who shared her non-remote sensing view on my thesis.

Last but not least, my thanks go to my family who gave support outside of the scientific framework and my friends and fellow geography students how largely contributed to making my time at UZH a memorable one.

References

- BAN, Y. & YOUSIF, O. 2016. Change Detection Techniques: A Review. In: BAN, Y. (ed.) *Multitemporal Remote Sensing: Methods and Applications*. Cham: Springer International Publishing.
- CHICCO, D. & JURMAN, G. 2020. The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC Genomics*, 21, 6.
- CHICCO, D., TOTSCH, N. & JURMAN, G. 2021. The Matthews correlation coefficient (MCC) is more reliable than balanced accuracy, bookmaker informedness, and markedness in two-class confusion matrix evaluation. *BioData Min*, 14, 13.
- DABOV, K., FOI, A., KATKOVNIK, V. & EGIAZARIAN, K. 2007. Image Denoising by Sparse 3-D Transform-Domain Collaborative Filtering. *IEEE Transactions on Image Processing*, 16, 2080-2095.
- DURIEUX, A. M., CALEF, M., ARKO, S., CHARTRAND, R., KONTGIS, C., KEISLER, R. & WARREN, M. 2019. *Monitoring forest disturbance using change detection on synthetic aperture radar imagery*, SPIE.
- FOODY, G. M. 2020. Explaining the unsuitability of the kappa coefficient in the assessment and comparison of the accuracy of thematic maps obtained by image classification. *Remote Sensing of Environment*, 239.
- GONG, M., LI, Y., JIAO, L., JIA, M. & SU, L. 2014. SAR change detection based on intensity and texture changes. *ISPRS Journal of Photogrammetry and Remote Sensing*, 93, 123-135.
- GUO, Z., LIU, H., PANG, L., FANG, L. & DOU, W. 2021. DBSCAN-based point cloud extraction for Tomographic synthetic aperture radar (TomoSAR) three-dimensional (3D) building reconstruction. *International Journal of Remote Sensing*, 42, 2327-2349.

- HACHICHA, S. & CHAABANE, F. 2014. On the SAR change detection review and optimal decision. *International Journal of Remote Sensing*, 35, 1693-1714.
- HECHELTJEN, A., THONFELD, F. & MENZ, G. 2014. Recent Advances in Remote Sensing Change Detection – A Review. *In: MANAKOS, I. & BRAUN, M. (eds.) Land Use and Land Cover Mapping in Europe: Practices & Trends*. Dordrecht: Springer Netherlands.
- HORN, R., NOTTENSTEINER, A. & SCHEIBER, R. 2008. DLR's advanced airborne SAR system onboard DO228. *7th European Conference on Synthetic Aperture Radar (EUSAR)*. VDE.
- HRIPCSAK, G. & ROTHSCHILD, A. S. 2005. Agreement, the f-measure, and reliability in information retrieval. *J Am Med Inform Assoc*, 12, 296-8.
- HUGUES, T., GOULETTE, F., DESCHAUD, J.-E., MARCOTEGUI, B. & LEGALL, Y. Semantic classification of 3D point clouds with multiscale spherical neighborhoods. 2018 International conference on 3D vision (3DV), 2018. IEEE, 390-398.
- HUTTNER, S., BRUSE, M., DOSTAL, P. & KATZSCHNER, A. Strategies for mitigating thermal heat stress in central European cities: The project Klimes. The Seventh International Conference on Urban Climate ICUC-7, 2009.
- KLETTE, R. & ROSENFELD, A. 2004. *Digital Geometry – Geometric Methods for Digital Picture Analysis*, San Francisco, Morgan Kaufmann.
- KROESE, D. P. & RUBINSTEIN, R. Y. 2012. Monte carlo methods. *Wiley Interdisciplinary Reviews: Computational Statistics*, 4, 48-58.
- LEBRUN, M. 2012. An analysis and implementation of the BM3D image denoising method. *Image Processing On Line*, 2012, 175-213.

- LI, Z. & BETHEL, J. 2008. Image coregistration in SAR interferometry. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 37, 433-438.
- LILLESAND, T., KIEFER, R. W. & CHIPMAN, J. 2015. *Remote Sensing and Image Interpretation, 7th Edition*, New York, John Wiley and Sons.
- LIN, Y., HONG, W., TAN, W. X., WANG, Y. P. & XIANG, M. S. 2012. Airborne Circular Sar Imaging: Results at P-Band. *2012 Ieee International Geoscience and Remote Sensing Symposium (Igarss)*, 5594-5597.
- MATLAB 2021. R2021b Update 3. (9.11.0.1873467) ed. Natick, Massachusetts: The MathWorks Inc.
- MENDEZ DOMINGUEZ, E., FRILOUD, M., SMALL, D. & HENKE, D. Range Adaptive Processing and Multisquint Processing Mode for SAR Image Change Detection. EUSAR 2018; 12th European Conference on Synthetic Aperture Radar, 4-7 June 2018 2018a. 1-6.
- MENDEZ DOMINGUEZ, E., MAGNARD, C., MEIER, E., SMALL, D., SCHAEPMAN, M. E. & HENKE, D. 2019. A Back-Projection Tomographic Framework for VHR SAR Image Change Detection. *IEEE Transactions on Geoscience and Remote Sensing*, 57, 4470-4484.
- MENDEZ DOMINGUEZ, E., MEIER, E., SMALL, D., SCHAEPMAN, M. E., BRUZZONE, L. & HENKE, D. 2018b. A Multisquint Framework for Change Detection in High-Resolution Multitemporal SAR Images. *IEEE Transactions on Geoscience and Remote Sensing*, 56, 3611-3623.
- MENDEZ DOMINGUEZ, E., SMALL, D. & HENKE, D. 2022. Back-Projection Tomography with Polarimetric HoloSAR data and Sub-aperture Processing. *International Geoscience and Remote Sensing Symposium IGARSS*. Kuala Lumpur: IEEE.

- MÉNDEZ DOMÍNGUEZ, E., SMALL, D. & HENKE, D. Synthetic Aperture Radar Tomography for Change Detection Applications. 2019 Joint Urban Remote Sensing Event (JURSE), 22-24 May 2019 2019. 1-4.
- MOULDEN, B., KINGDOM, F. & GATLEY, L. F. 1990. The Standard Deviation of Luminance as a Metric for Contrast in Random-Dot Images. *Perception*, 19, 79-101.
- PATEL, V. M., EASLEY, G. R., HEALY, D. M. & CHELLAPPA, R. 2010. Compressed Synthetic Aperture Radar. *IEEE Journal of Selected Topics in Signal Processing*, 4, 244-254.
- PRICE, B., KIENAST, F., SEIDL, I., GINZLER, C., VERBURG, P. H. & BOLLIGER, J. 2015. Future landscapes of Switzerland: Risk areas for urbanisation and land abandonment. *Applied Geography*, 57, 32-41.
- QUIN, G., PINEL-PUYSSÉGUR, B., NICOLAS, J. M. & LOREAUX, P. 2014. MIMOSA: An Automatic Change Detection Method for SAR Time Series. *IEEE Transactions on Geoscience and Remote Sensing*, 52, 5349-5363.
- RICHARDS, J. A. 2022. *Remote Sensing Digital Image Analysis*, Cham, Springer.
- SAID, K. A. M., JAMBEK, A. B. & SULAIMAN, N. 2016. A Study of Image Processing Using Morphological Opening and Closing Processes. *International Journal of Control Theory and Applications*, 9, 15 – 21.
- SUTTON, C. & MCCALLUM, A. 2012. An Introduction to Conditional Random Fields. *Foundations and Trends® in Machine Learning*, 4, 267-373.
- VIEIRA, S. M., KAYMAK, U. & SOUSA, J. M. C. Cohen's kappa coefficient as a performance measure for feature selection. International Conference on Fuzzy Systems, 18-23 July 2010 2010. 1-8.

- WALLACH, H. M. 2004. Conditional random fields: An introduction. *Technical Reports (CIS)*, 22.
- WARRENS, M. J. 2015. Five Ways to Look at Cohen's Kappa. *Journal of Psychology & Psychotherapy*, 05.
- WOODHOUSE, I. H. 2006. *Introduction to Microwave Remote Sensing*, Boca Raton, CRC Press.
- YI-MIN, H. & SHU-XIN, D. Weighted support vector machine for classification with uneven training class sizes. 2005 International Conference on Machine Learning and Cybernetics, 18-21 Aug. 2005 2005. 4365-4369 Vol. 7.
- YU, X., HYYPPÄ, J., KARJALAINEN, M., NURMINEN, K., KARILA, K., VASTARANTA, M., KANKARE, V., KAARTINEN, H., HOLOPAINEN, M., HONKAVAARA, E., KUKKO, A., JAAKKOLA, A., LIANG, X., WANG, Y., HYYPPÄ, H. & KATOH, M. 2015. Comparison of Laser and Stereo Optical, SAR and InSAR Point Clouds from Air- and Space-Borne Sources in the Retrieval of Forest Inventory Attributes. *Remote Sensing*, 7, 15933-15954.
- ZHANG, J., CUI, M. & WANG, B. SAR Image Change Detection Method Based on Neural-CRF Structure. 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS, 11-16 July 2021 2021. 3797-3800.

Appendix

A. Accuracy assessment – baseline

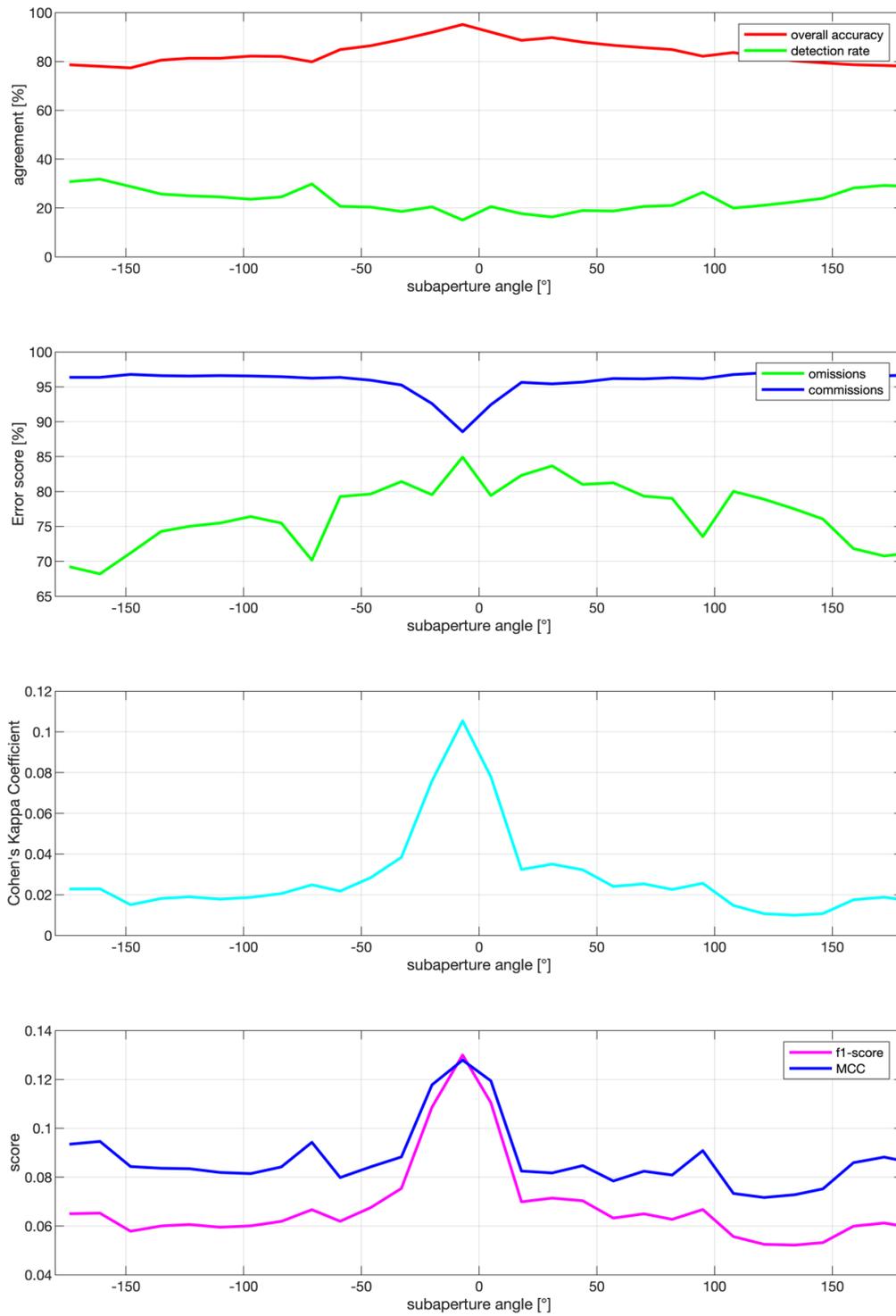


Figure 18: Accuracy assessment of the baseline method with all subapertures of 0118 vs. 0111 (southern linear flight of Memmingen).

Table 4: Detailed accuracy statistics for 0118 vs. 0110 evaluated on changes using the baseline method. Confusion matrix counts in number of pixels.

Approximative angle [°]	Subaperture First pulse	True positives	False negatives	False positives	True negatives	Overall accuracy [%]	F1-score
-2	1000	6376	10317	12164	662785	96.75	0.362
10	10000	6670	10023	31629	643320	93.98	0.243
23	19000	5908	10785	72187	602762	88.00	0.125
36	28000	5332	11361	58390	616559	89.92	0.133
49	37000	5518	11175	70337	604612	88.21	0.119
62	46000	5693	11000	82044	592905	86.55	0.109
74	55000	5839	10854	90103	584846	85.40	0.104
87	64000	6435	10258	105142	569807	83.32	0.100
100	73000	6697	9996	136452	538497	78.83	0.084
113	82000	6352	10341	118040	556909	81.44	0.090
126	91000	6193	10500	121268	553681	80.95	0.086
138	100000	6864	9829	131047	543902	79.63	0.089
151	109000	7680	9013	138061	536888	78.74	0.095
164	118000	7213	9480	138915	536034	78.54	0.089
177	127000	6304	10389	147372	527577	77.19	0.074
190	136000	5298	11395	131237	543712	79.38	0.069
203	145000	5179	11514	112037	562912	82.14	0.077
215	154000	5257	11436	105954	568995	83.03	0.082
228	163000	5516	11177	101629	573320	83.69	0.089
241	172000	5069	11624	96367	578582	84.39	0.086
254	181000	4902	11791	85338	589611	85.96	0.092
267	190000	5098	11595	90728	584221	85.21	0.091
279	199000	4557	12136	79337	595612	86.77	0.091
292	208000	4875	11818	67511	607438	88.53	0.109
305	217000	5019	11674	55596	619353	90.27	0.130
318	226000	5270	11423	41730	633219	92.31	0.165
331	235000	5869	10824	23459	651490	95.04	0.255
343	244000	6525	10168	14276	660673	96.47	0.348
356	253000	5429	11264	38556	636393	92.80	0.179

Table 5: Detailed accuracy statistics for 0118 vs. 0111 evaluated on changes using the baseline method. Confusion matrix counts in number of pixels.

Approximative angle [°]	Subaperture First pulse	True positives	False negatives	False positives	True negatives	Overall accuracy [%]	F1-score
-174	1000	5137	11556	136202	538747	78.64	0.065
-161	10000	5308	11385	140750	534199	78.00	0.065
-148	19000	4807	11886	144705	530244	77.36	0.058
-135	28000	4292	12401	122067	552882	80.56	0.060
-123	37000	4170	12523	116783	558166	81.30	0.061
-110	46000	4092	12601	116778	558171	81.29	0.059
-97	55000	3937	12756	110501	564448	82.18	0.060
-84	64000	4093	12600	111522	563427	82.05	0.062
-71	73000	4980	11713	127727	547222	79.84	0.067
-59	82000	3459	13234	91545	583404	84.85	0.062
-46	91000	3397	13296	80509	594440	86.44	0.068
-33	100000	3100	13593	62460	612489	89.00	0.075
-20	109000	3415	13278	42752	632197	91.90	0.109
-7	118000	2515	14178	19486	655463	95.13	0.130
5	127000	3433	13260	42015	632934	92.01	0.110
18	136000	2951	13742	64798	610151	88.64	0.070
31	145000	2725	13968	56902	618047	89.75	0.071
44	154000	3169	13524	70269	604680	87.88	0.070
57	163000	3129	13564	79141	595808	86.60	0.063
70	172000	3449	13244	86026	588923	85.65	0.065
82	181000	3504	13189	91549	583400	84.86	0.063
95	190000	4417	12276	111271	563678	82.14	0.067
108	199000	3334	13359	99780	575169	83.64	0.056
121	208000	3523	13170	114011	560938	81.61	0.052
134	217000	3757	12936	123566	551383	80.26	0.052
146	226000	3996	12697	129582	545367	79.43	0.053
159	235000	4704	11989	135636	539313	78.66	0.060
172	244000	4880	11813	137885	537064	78.36	0.061
185	253000	4796	11897	140483	534466	77.97	0.059

B. Accuracy Assessment – Find Pairs

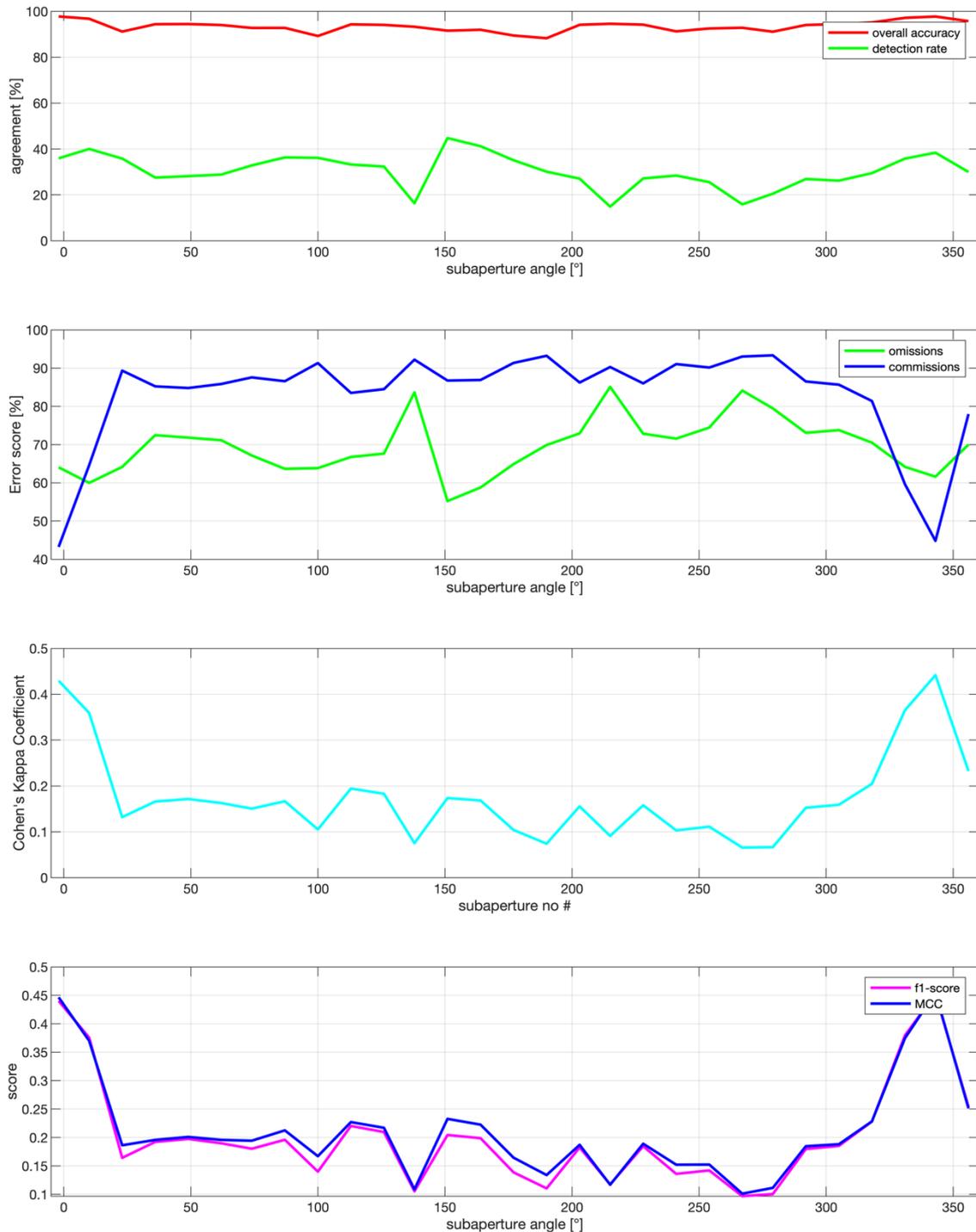


Figure 19: Accuracy assessment after the application of the Find Pairs algorithm showing detection rate, commissions and omissions as well as the kappa coefficient, F1-score and Matthew's correlation coefficient. Data of all subapertures from flight 0118 against 0110 (northern linear flight of Memmingen).

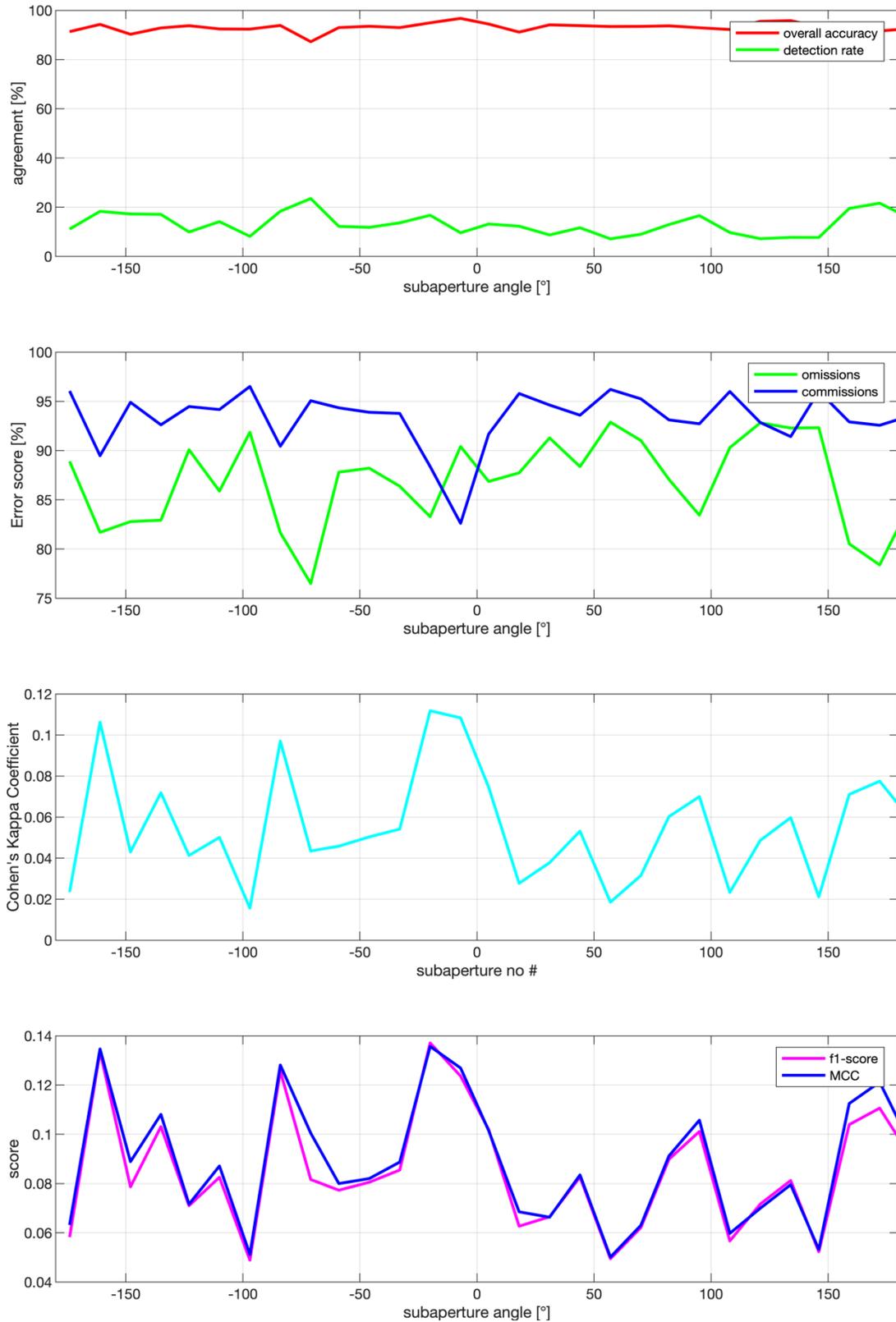


Figure 20: Accuracy assessment after application of the find pairs algorithm. All subapertures of 0118 vs. 0111 (southern linear flight of Memmingen)

C. Accuracy assessment – comparison

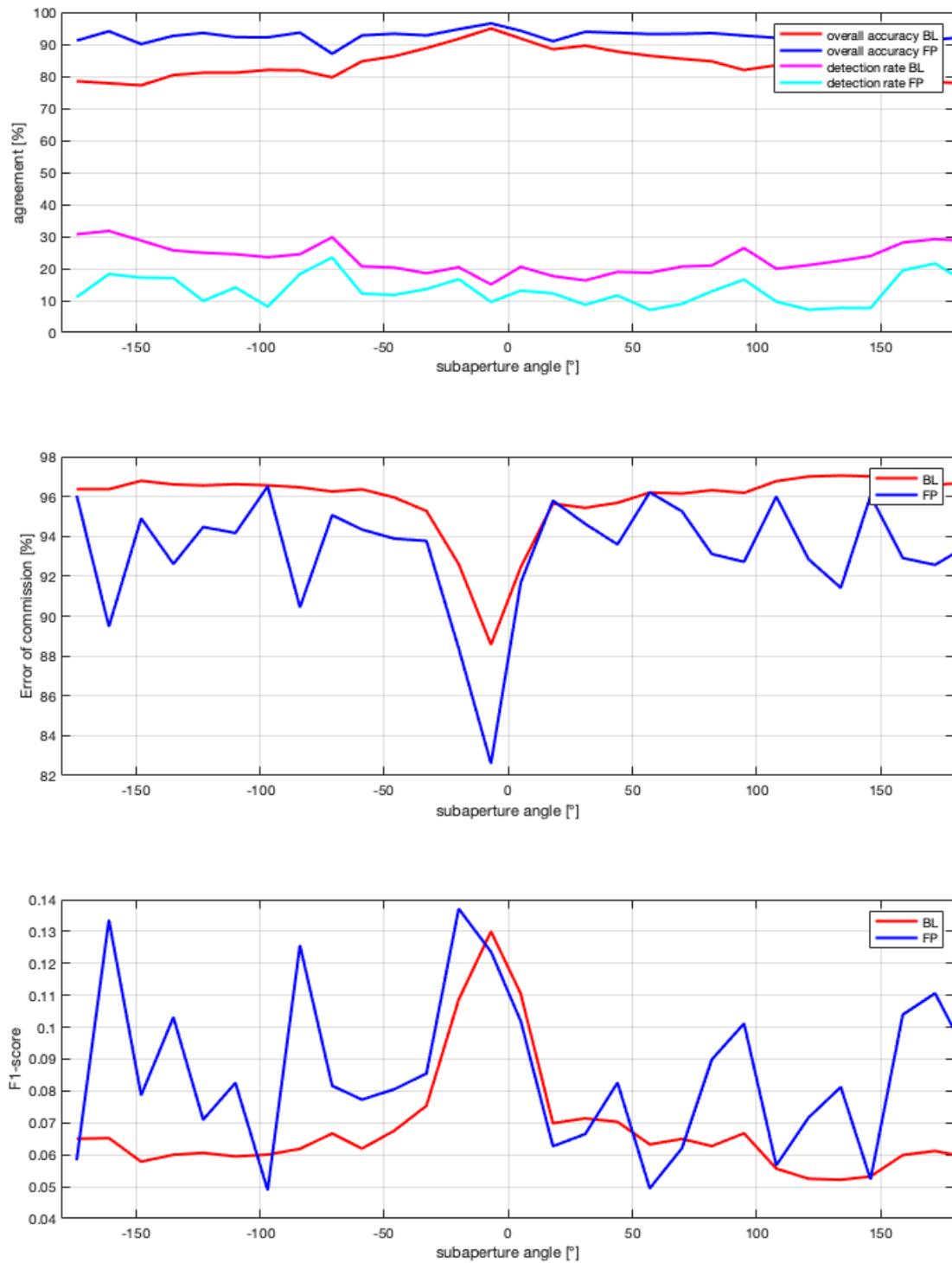


Figure 21: Comparison of the classification performance of the baseline method (BL) to Find Pairs (FP). Subapertures of dataset 0118 as reference images and 0111 as test image.

Table 6: Detailed accuracy statistics for 0118 vs. 0110 evaluated on changes using Find Pairs. Confusion matrix counts in number of pixels.

Approximative angle [°]	Subaperture First pulse	True positives	False negatives	False positives	True negatives	Overall accuracy [%]	F1-score
-2	1000	6000	10693	4570	670379	97.79	0.440
10	10000	6680	10013	12186	662763	96.79	0.376
23	19000	5982	10711	50167	624782	91.20	0.164
36	28000	4597	12096	26540	648409	94.41	0.192
49	37000	4703	11990	26229	648720	94.47	0.198
62	46000	4818	11875	29250	645699	94.05	0.190
74	55000	5484	11209	38673	636276	92.79	0.180
87	64000	6067	10626	39143	635806	92.80	0.196
100	73000	6035	10658	63500	611449	89.28	0.140
113	82000	5549	11144	28124	646825	94.32	0.220
126	91000	5401	11292	29442	645507	94.11	0.210
138	100000	2730	13963	32355	642594	93.30	0.105
151	109000	7475	9218	48966	625983	91.59	0.204
164	118000	6881	9812	45648	629301	91.98	0.199
177	127000	5860	10833	61968	612981	89.47	0.139
190	136000	5027	11666	69224	605725	88.30	0.111
203	145000	4519	12174	28316	646633	94.15	0.182
215	154000	2487	14206	23125	651824	94.60	0.118
228	163000	4534	12159	27881	647068	94.21	0.185
241	172000	4747	11946	48356	626593	91.28	0.136
254	181000	4266	12427	39076	635873	92.55	0.142
267	190000	2649	14044	35367	639582	92.86	0.097
279	199000	3426	13267	48066	626883	91.13	0.100
292	208000	4492	12201	28805	646144	94.07	0.180
305	217000	4375	12318	26166	648783	94.44	0.185
318	226000	4922	11771	21542	653407	95.18	0.228
331	235000	5979	10714	8833	666116	97.17	0.380
343	244000	6409	10284	5206	669743	97.76	0.453
356	253000	5006	11687	17735	657214	95.75	0.254

Table 7: Detailed accuracy statistics for 0118 vs. 0111 evaluated on changes using Find Pairs. Confusion matrix counts in number of pixels.

Approximative angle [°]	Subaperture First pulse	True positives	False negatives	False positives	True negatives	Overall accuracy [%]	F1-score
-174	1000	1852	14841	45059	629890	91.34	0.058
-161	10000	3056	13637	25989	648960	94.27	0.134
-148	19000	2874	13819	53497	621452	90.27	0.079
-135	28000	2850	13843	35742	639207	92.83	0.103
-123	37000	1655	15038	28269	646680	93.74	0.071
-110	46000	2356	14337	38042	636907	92.43	0.083
-97	55000	1359	15334	37551	637398	92.35	0.049
-84	64000	3061	13632	28967	645982	93.84	0.126
-71	73000	3924	12769	75518	599431	87.24	0.082
-59	82000	2035	14658	33934	641015	92.97	0.077
-46	91000	1969	14724	30238	644711	93.50	0.081
-33	100000	2274	14419	34239	640710	92.96	0.085
-20	109000	2790	13903	21218	653731	94.92	0.137
-7	118000	1601	15092	7604	667345	96.72	0.124
5	127000	2193	14500	24155	650794	94.41	0.102
18	136000	2048	14645	46613	628336	91.14	0.063
31	145000	1455	15238	25612	649337	94.09	0.066
44	154000	1941	14752	28355	646594	93.77	0.083
57	163000	1188	15505	30158	644791	93.40	0.049
70	172000	1498	15195	30059	644890	93.46	0.062
82	181000	2158	14535	29189	645760	93.68	0.090
95	190000	2765	13928	35219	639730	92.89	0.101
108	199000	1620	15073	38848	636101	92.20	0.057
121	208000	1199	15494	15574	659375	95.51	0.072
134	217000	1287	15406	13710	661239	95.79	0.081
146	226000	1282	15411	31046	643903	93.28	0.052
159	235000	3251	13442	42616	632333	91.89	0.104
172	244000	3609	13084	44948	630001	91.61	0.111
185	253000	2582	14111	37333	637616	92.56	0.091

D. Results of an independent, suboptimal dataset

Acquisition flight 0418 builds a set of data that is independent from the other data despite being recorded with the same equipment. In order to test the algorithm developed in this thesis, it was applied on this data set to check its performance on data which it has not been optimized on.

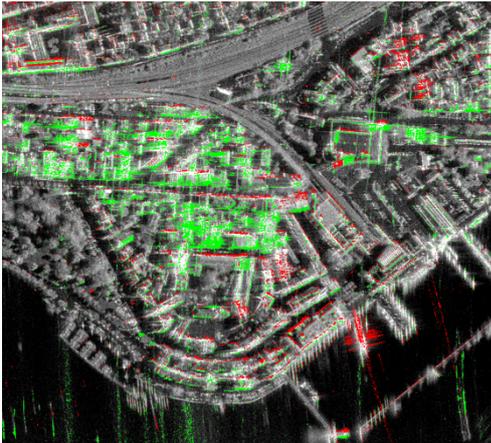
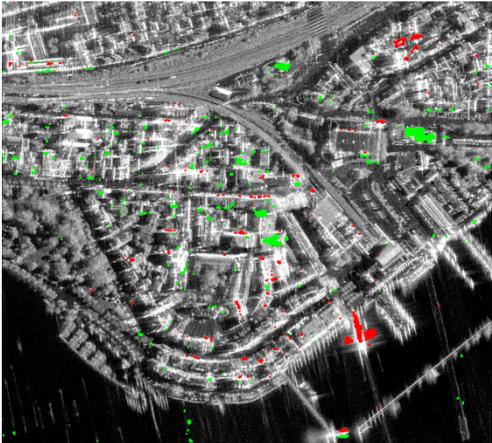
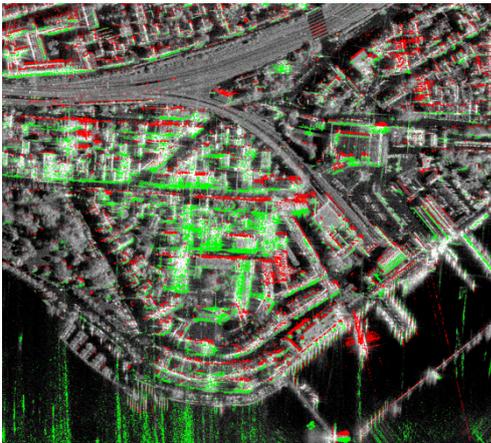
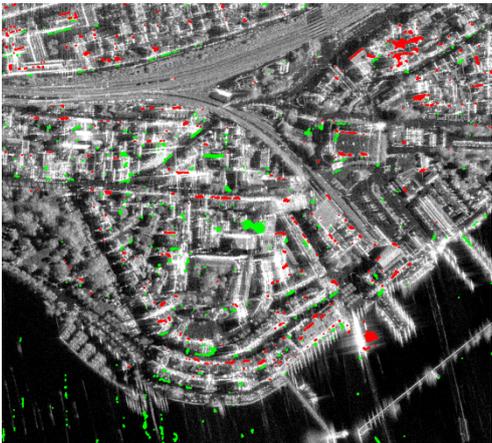
Baseline Method	Find Pairs	Subaperture
		224 vs. 226 ~2°
		224 vs. 232 ~8°
		224 vs. 240 ~ 17°

Figure 22: Comparison of the algorithm performance of the validation dataset. Data within the circle are used for both reference and test image. Change detection is applied on different between images of different viewing angles of the same circle.

Examples of the results can be viewed in Figure 22. At a 2°-deviation of the illumination angle, the baseline method shows some lines of positive changes. They are formed as thin lines adjoining edges of buildings. Negative changes occur to a much lesser degree in comparison. Using Find Pairs eliminates a large part of this. Only the top right corner still shows some positive changes and some more compact forms of negative change have been retained.

The second example with a deviation of 8° shows an even larger number of positive changes which still outnumber the negative changes by a large scale. It appears that the positive changes are concentrated around the upper center of the image while the negative are more prevalent in the lower part of the image. A presumably tall structure at the shore, possibly some sort of pole or tower, shows a strong backscatter in the first example and is covered by negative change in the second example. Find Pairs again eliminates many of the positive changes seen in the baseline image and also reduces the occurrence of negative changes. It seems to aggregate at least some of the changes and some crumbling occurs, similar to the phenomenon observed in some examples of Figure 15. There are few larger elements of change which mostly are preserved through Find Pairs.

The third example with 17° of angular distance shows well developed elements of detected change. They occur as elongated structures around and between buildings. While the positive change class is more dominant, negative changes occur more often than in the previous two examples. Visually judging, Find Pairs reduces the amount of changes again while it also changes the structure of the changes to more compact, rounded forms and mostly eliminates the elongated forms.

It is noteworthy that in all examples, most changes appear as accumulation of thin lines mostly featuring only few pixels in width. After the treatment with find changes, more aggregated, compact forms of change elements are prevalent.

While this dataset does not contain any true changes, the behavior of Find Pairs could be evaluated to some degree on an independent dataset. It shows that Find Pairs conducts a rather significant removal of false alarms, even though not all of them can be attributed to pairs. Here, it can be assumed that the morphological preprocessing as well as the effect described in Section 4.2.3 play an important role.

Personal Declaration

I hereby declare that the submitted Thesis is the result of my own, independent work.
All external sources are explicitly acknowledged in the Thesis.

Zurich, September 30th, 2022

Signed

A handwritten signature in black ink. The signature consists of a large, stylized 'L' that loops around and underlines the word 'Ritter', which is written in a cursive, handwritten style.

Lukas Ritter