Universität Augsburg

Fakultät für Angewandte Informatik

Institut für Geographie

Region based building footprint extraction and change detection for urban areas

Betreut durch:

Dr. Jiaojiao Tian

PD Dr. Andreas Phillipp

Metzlaff, Lukas

Matrikelnummer: 1086596

Master Geoinformatik 4. Semester

Abgabetermin: 11.05.2015

Table of contents

Table of figures

Table of tables

1 Introduction

Up-to-date geodata is important for a lot of applications like disaster management, or urban planning (Gamba and Houshmand, 2000). Thereby, the availability and quality of remote sensing based geodata increased largely in the last decades. Making use of “Big Data” has therefore become an important research topic.

Especially the automatic extraction of building footprints and the detection of building changes has thereby a high scientific value and therefore many methods were proposed. These differ on the one side dependent on the used data. While some are using just photographic data, others are using height data in addition. In previous research there are thereby two main sources for the height data used. One is LIDAR (Light detection and ranging) - data, which has a very high accuracy. The other source is stereo imagery, were a DSM is calculated out of satellite stereo data. These stereo data based DSM are not as accurate as the LIDAR based DSM’s. However, they got the advantage of a better availability for large regions.

For change detection, a major distinction can be seen between pixel and object based approaches. While pixel based approaches determining the building change for each pixel independently, object based approaches are determining the change based on groups of pixels. However, how to obtain accurate regions for building changes is still an open topic.

Therefore, in thesis both building footprint detection and change detection is performed, based on satellite stereo data. Thereby a comparison is made between a just pixel based approach for change detection and methods that are object based.

At first the state of art in terms of change detection and building footprint extraction is outlined. After that, the background for some of the basic functions is clarified. Next, the method explained in detail, before the results are described and the method is discussed. In the end there is a conclusion.

2 State of art

2.1 Building footprint extraction

For urban planning, map updates, insurances, hazard prevention and real estate monitoring it is important to have actual data about the buildings in an area. Especially in urban areas, where the façades, the structures and alignment of buildings change fast, it is important to update data efficiently and reliably; therefore, there is a growing interest in the development of tools and sensors used for the detection of buildings – with reduced working times and reasonable costs (Gamba and Houshmand, 2000). For this task,

multiple approaches are proposed that can be categorized into image, LIDAR and stereo satellite imagery based.

2.1.1 Image based building footprint extraction

Because of the fact that available height data does not always exist, it is vital to have methods for building detection on the basis of multispectral and panchromatic images (Shorter and Kasparis, 2009).

To achieve this, the above-mentioned multiple approaches were proposed; they are subdivided into automatic and supervised ones. Supervised means that there needs to be a user supervised training phase in the beginning while automatic approaches work independently from the scene and the user. Furthermore, a distinction between area and feature based approaches can be made. With feature based approaches, geometric information (Corners, Lines) is examined while in area based approaches multispectral/panchromatic information is utilized (Shorter and Kasparis, 2009).

Lee et al. (2003) stated that building extraction is achieved by using high resolution (1 m or better) multispectral and panchromatic data from Ikonos. Thereby, at first supervised land cover classification was performed that is based on the multispectral data with the focus on detecting building roofs. These classification results were further refined with segmentation based on the panchromatic image.

Ünsalan and Boyer (2005) also used IKONOS imagery to detect houses and street networks. Contrary to the previous method they used an automatic approach that calculates an index for human activity. Based on this approach, clustering and classification was performed and initial building candidate areas were extracted. These were further refined based on their geometric properties.

An as well automatic combined areal and feature based approach was presented by (Shorter and Kasparis, 2009). They used the multispectral information for colour based segmentation by excluding vegetation and shadow regions. Further they used geometric information by assuming that buildings are areas with a convex hull.

As in the before mentioned approaches rather more general assumptions are used, other automatic approaches define more specific rules for buildings. Lefevre et al. (2007) for example specified buildings as rectangular while Muller and Zaum (2005) declared building roofs to be red. Therefore, these approaches depend on the scene.

To summarize the above mentioned information: the image based footprint extraction usually starts with some kind of pixel based classification (roof, street, water, vegetation, etc.). Through clustering/segmentation the accuracy of the classification is improved subsequently. Building candidate areas are determined whereby unneeded classes (shadows etc.) are excluded. Finally, buildings are extracted either based on the classification by itself or in combination with the geometric segment attributes.

2.1.2 LIDAR based building footprint extraction

LIDAR (LIght Detection And Ranging) data provides DSM's with a very good accuracy and is therefore ideal to apply for footprint extraction. A peculiarity of LIDAR is thereby that these systems are able to register two echoes of a laser beam. The first pulse is a reflection of the vegetation while the second pulse corresponds to the lowest reflection point (soil, etc.). Due to this, it is possible to detect the ground level, which can be useful for the DTM (digital terrain model) generation. For the extraction of the DTM, further several different morphological filtering techniques can be used. The DTM thereby represents the terrain surface (DSM without vegetation and buildings) (Rottensteiner et al. 2003).

Some of the approaches were based on the comparison of the DSM with the DTM. By doing so, the normalised DSM (nDSM) representing the height of all objects (vegetation, buildings) above ground level was calculated. As the focus is on building extraction, the vegetation areas need to be eliminated from the nDSM. This can be with the LIDAR data itself because the laser does not just deliver height information, but also information on the physical properties of the materials. Thus vegetation can be detected and excluded (Gamba and Houshmand, 2000). Other approaches used additional multispectral data to calculate the NDVI (normalized difference vegetation index). Based on this data, it was possible to detect vegetation areas and exclude them as well (Rottensteiner et al. 2003). Through the refined nDSM initial building candidate regions were extracted by using height based thresholds – which were further refined in following steps.

Other approaches (Zhou and Troy, 2008) did not rely that much on the nDSM generation. In the classification procedure they utilized the LIDAR as one among many features. Zhou and Troy (2008) thereby segmented the image at first. Furthermore, a combination of the LIDAR and multispectral data served to subdivide the segments into buildings, pavements, streets and different types of vegetation. The building footprint was thereby given by the building class.

2.1.3 Satellite stereo image based

As well as the LIDAR based approaches, the satellite stereo image based approaches also use DSM's for the footprint extraction. Due to the increasing quality of satellite image data, the global coverage and the lower costs, satellite based stereo data has become more popular in the recent years (Tian et al. 2014; San and Turker, 2006). The DSM's data generated from the satellite stereo are, however, not as precise as those from LIDAR (Lu et al. 2002).

When using stereo image data, the DSM needs to be at first calculated by stereo image matching algorithms. After that, the methods are pretty similar to the LIDAR based approaches.

In Lu et al. (2002) classification was performed based on multispectral data. The multispectral data was used together with a DTM to identify building candidate regions. By using the NDVI, vegetation was excluded from the candidate regions. Out of these candidate areas the building footprint was extracted. Thereby, a shape modelling algorithm was used for the precise delineation of the buildings.

A similar approach was presented by (San and Turker 2006). They also used a type of classification based on additional multispectral data. Based on the classification and a generated nDSM building candidate regions were identified. In addition these candidate areas were compared with an already existing building database. By doing so, buildings missing out the footprint were deleted from the database. Candidate areas above a certain size not represented in the database were assumed to be new buildings and were updated. Thereby, the shape of the neighbouring buildings was examined and it was assumed that the new construction generated at the centroid position of the building candidate area was expected to have the same shape.

As shown with the mentioned examples, building footprint extraction based on satellite stereo images is usually based on the generation of building candidate areas calculated with the help of an nDSM and multispectral data. The method presented in this master thesis refers to a similar approach. Thereby, an nDSM is utilized for the generation of the initial candidate regions. Besides that, multispectral data shadows are calculated which are excluded from the candidate region. Finally, height based local thresholds are used to determine the building segments within a candidate region.

2.2 Change detection

Change detection procedures detect change features by comparing data from two different times. These methods can be categorized by the used data into 2D and 3D approaches, where 3D stands for the use of height data. Further pixel and object based approaches can be distinguished (Tian, 2013). Thereby, in the pixel based approaches change features are compared pixel wise while object based methods compare these features on segment/object level (Tian et al. 2013).

2.2.1 Pixel based change detection

One of the most basic pixel based approaches is thereby *Image differencing* where two images are simply subtracted. This can be done directly on panchromatic images or for change features like the NDVI for instance (Hayes and Sader, 2001).

For the comparison of multiple features, as it is the case when using a multispectral data directly, *change vector analysis* (CVA) can be used. Thereby, a change vector is calculated containing the information about the differences between the features of two

pixels. The magnitude of the vector thereby provides information about the change probability while the vector direction can be used to identify the type of change (Chen et al. 2003).

Another common approach is the *post-classification comparison*. Thereby, two independently generated classification maps are compared (Liu and Zhou, 2004). This can also be performed with two building footprints, as applied for one of the change detection methods in this master thesis.

However, pixel based methods can also be more complex. *Statistical based similarity measures* for instance calculate the theoretical similarity based on mutual information in both images. By doing so, a change indicator image can be generated out of multiple input features (Gueguen et al., 2011). Different from that, *multivariate alteration detection* can as well be used for change detection with multiple features. Thereby, it is not just focused on change, but also on 'no change' indicators. This helps for the identification of real changes (Nielsen et al., 1998).

2.2.2 Object based change detection

One problem of the pixel based approaches is the salt-and-pepper noise. By using object based approaches, changes can be detected more accurately and with a fewer rate of false alarms (Blaschke, 2010).

Therefore, it is necessary to gain information on meaningful homogeneous regions by using already levied data from GIS databases (Durieux et al., 2008). Thereby, either the defined object are directly used as the areas for change detection or based on the data, training for a classification can be performed (Walter, 2004).

Another approach to get homogeneous regions is using segmentation methods. Thereby, based on different input features like panchromatic- , height-, and multispectral images, clustering can be performed. These approaches got the advantage being not limited to just the already available data. Thus, approaches just based on images are preferred in recent studies (Tian, 2013).

The accuracy of object based approaches proved to be better than pixel-based methods in recent studies (Blaschke, 2010). However, this conclusion is built on a hypothesis that proper region maps are available. The prefect case would be each building is presented as a single region in the region map. Therefore, this thesis focuses on improving the segmentation results, and further uses them for building footprint extraction and change detection.

3 Background

Subsequently, the background for some of the used functions in the method is explained.

3.1 Image Segmentation

The goal of segmentation is to group pixels based on the similarity of their attributes. Segmentation thereby consists of two tasks. One is to recognize objects, while the second is to delineate these ones (de Miranda et al., 2010). In order to accomplish these tasks many methods were introduced. As some of these methods work independently, others depend on user input.

Especially for the recognition of objects, user support can help a lot to improve the segmentation results. Often these approaches use markers for the object recognition (de Miranda et al., 2010; Ning et al. 2010). Thereby, the user needs to set object and background markers. Based on this method, merging is performed on a previous unsupervised segmentation. By doing so, the marked objects are merged to one segment, while the rest is assigned as background.

Despite the benefits of supervised segmentation, automatic segmentation methods are more suitable for building footprint detection and change detection. Some of the used methods are thereby level-set, watershed (Vincent and Soille, 1991) and mean-shift (Comaniciu and Meer, 2002), which will be used in this master thesis.

The mean-shift method is able to determine an unpredefined number of arbitrary shaped clusters in a feature space. This is a multidimensional space where the pixels from the input images are represented as points (see (a) for a 2D feature space) (Comaniciu and Meer, 2002). These points contain all the information from input images. Attributes used in the approach in this master thesis are panchromatic intensity, the x- and y- position of the pixel, height value and multispectral information.

The goal of the mean shift is to search and delineate clusters within the feature space, which are areas with a high point density. This is achieved by examining the neighbourhood of each data point within a predefined kernel shape and size. Within this kernel the mean value for all attributes is calculated. The kernel is subsequently shifted to the position of the calculated mean values. By doing so, the kernel shifts to the local highest density area, as shown with the black trajectories in Figure 3-1 (c). The density is displayed on the vertical axes. When the calculated mean values in the kernel stay the same, convergence is reached and the point with the local highest density is found.

The clusters (see Figure 3-1 (b)) are then given by the “basin of attraction” (Comaniciu and Meer, 2002, S.608) of a convergence point.

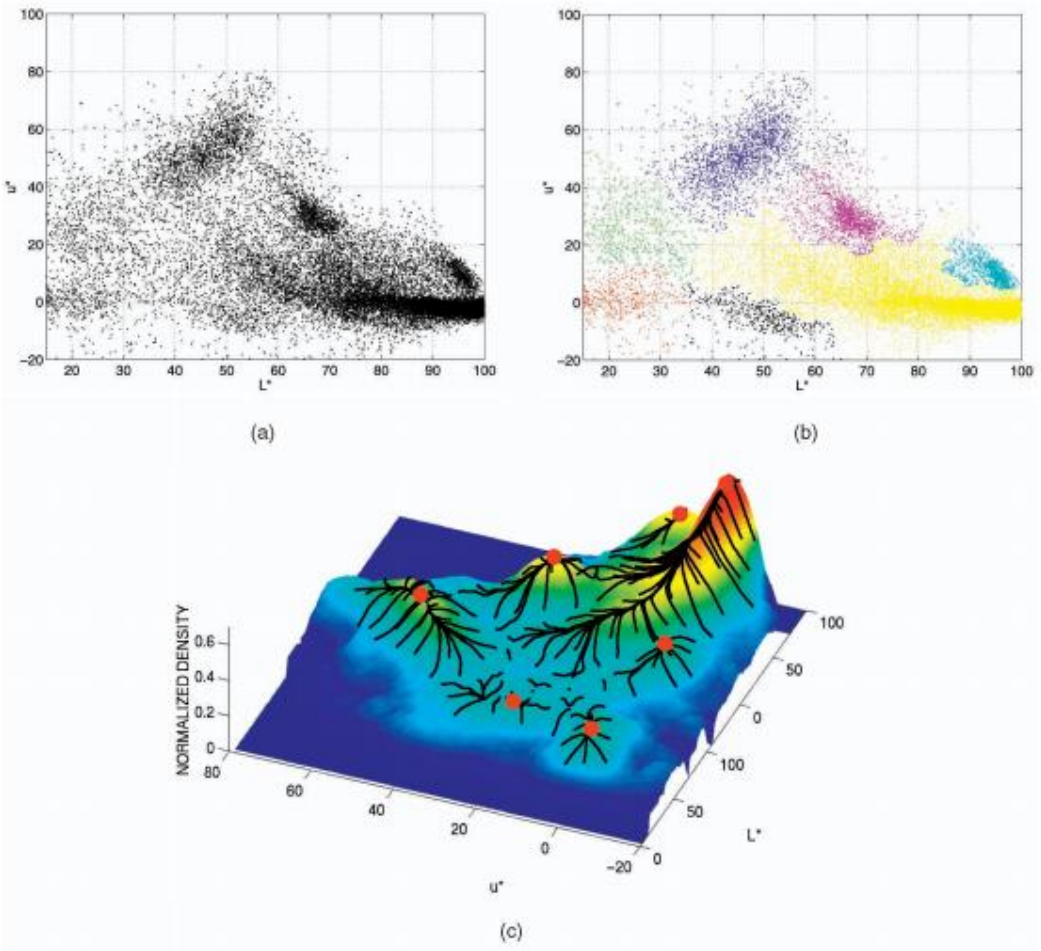

Figure 3-1: 2D feature space analysis. (a) 2D feature space of a colour image in L*u*v representation. (b) Resulting clusters. (c) Point density of the feature space; black lines are representing the trajectories of the mean shift procedure; red dots are marking the convergence points.

Source: Comaniciu and Meer, 2002, S.609.

3.2 Sobel operator

The sobel operator is used to calculate the gradient and emphasizes edges in an image. Thereby the image (A) is convolved $(*)$ with a horizontal (G_x) and a vertical (G_y) 3x3 filter, whereby the derivatives are calculated (Scharr, 2000).

(1)

$$G_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} * A \qquad G_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} * A$$

To calculate the magnitude of the gradient (G), there are two ways possible:

(2)

$$G = G_x + G_y$$

(3)

$$G = \sqrt{{G_x}^2 + {G_y}^2}$$

Thereby, in this work equation (2) is used for the calculation of the magnitude of the gradient, for it is faster to calculate.

3.3 Morphological filters

In this work two morphological filters are used. One of them is erosion. Thereby, the value of a pixel is set to the minimum value in a defined neighbourhood, which leads to an enlargement of small valued areas (Serra, 1994).

The other morphological filter used in this paper is dilation. Dilation is thereby the opposite of erosion. At that, pixel (i,j) is assigned the maximum value in a given neighbourhood. In doing so, areas with high values are enlarged (Serra, 1994).

The combination of erosion at first and a following dilation is called opening. At that, the erosion leads to a complete removal of small bright spots. However, bigger bright areas are kept and subsequently retained by the dilation. Opening is used in this method to separate building candidate areas from another (Serra, 1994).

The counterpart of opening is closing. At that, dilation is applied before erosion. In doing so, small dark spots are removed. Further, this also leads to a flatting of rough boundaries from bright areas. For that, closing is used to improve the shapes of the objects in the building footprint (Serra, 1994).

For nDSM generation morphological reconstruction is used. One of the morphological reconstruction methods is thereby the top-hat filter. Thus, at first morphological opening is performed on the DSM to generate the DTM (digital terrain model). Thereby, small structures with larger heights than their surroundings (buildings, vegetation) are eliminated. The DTM is then subtracted from the DSM. As a result, the underlying elevation is removed and only objects above ground level are represented in the nDSM (Tian and Reinartz, 2013).

4 Method

To improve the accuracy of building extraction and change detection, -the following explained method is used. Firstly a brief overview over the manner of functioning and the major components are described. And then, these components are explained in detail.

4.1 Overview

The input image data in the presented workflow are generally just from one date. The only exclusion thereby is the change detection in the end, where there are the results from both dates needed for the comparison.

The first step of this method is the digital surface model (DSM) refinement. As can be seen in Figure 4-1, this is based on the original DSM and the panchromatic image data. With the help of a gradient based method, the improvement of the DSM in terms of the set goals could be achieved. This refined DSM is an important input for further steps. It will be used in the segmentation procedure as well as the building footprint extraction, and has a large impact on both steps.

The Segmentation takes other than the refined DSM also the panchromatic and the pansharpened multispectral image as input. The multispectral channels are at the beginning reduced with a principal components analysis and after preprocessing of all the input features a mean shift segmentation is performed.

The obtained region map after segmentation is used in the next phase together with the refined DSM and the multispectral data for the building footprint extraction. For that purpose the shadows and the normalized DSM (nDSM) are calculated, as is shown in Figure 4-2. This is used than to obtain building location candidates and the resulting building footprint.

The change detection result could be obtained then either by comparing the building footprints of both dates, for which the whole procedure needs to be applied to data of both time points, or by using the method just once. Thereby a change indicator needs to be processed, which reveals the new buildings out of the gained building footprint.

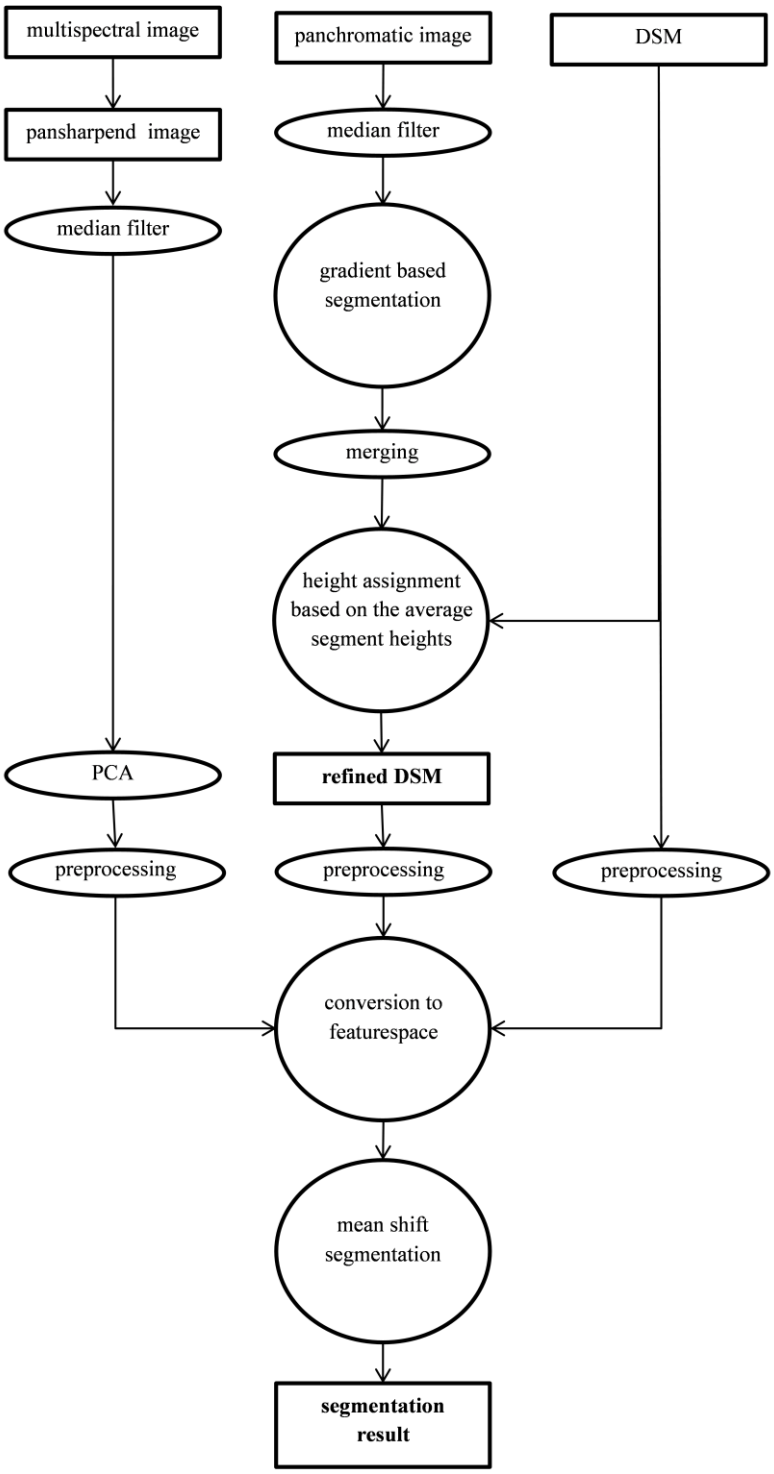

Figure 4-1: Workflow part 1 (calculation of the refined DSM and mean shift segmentation result)

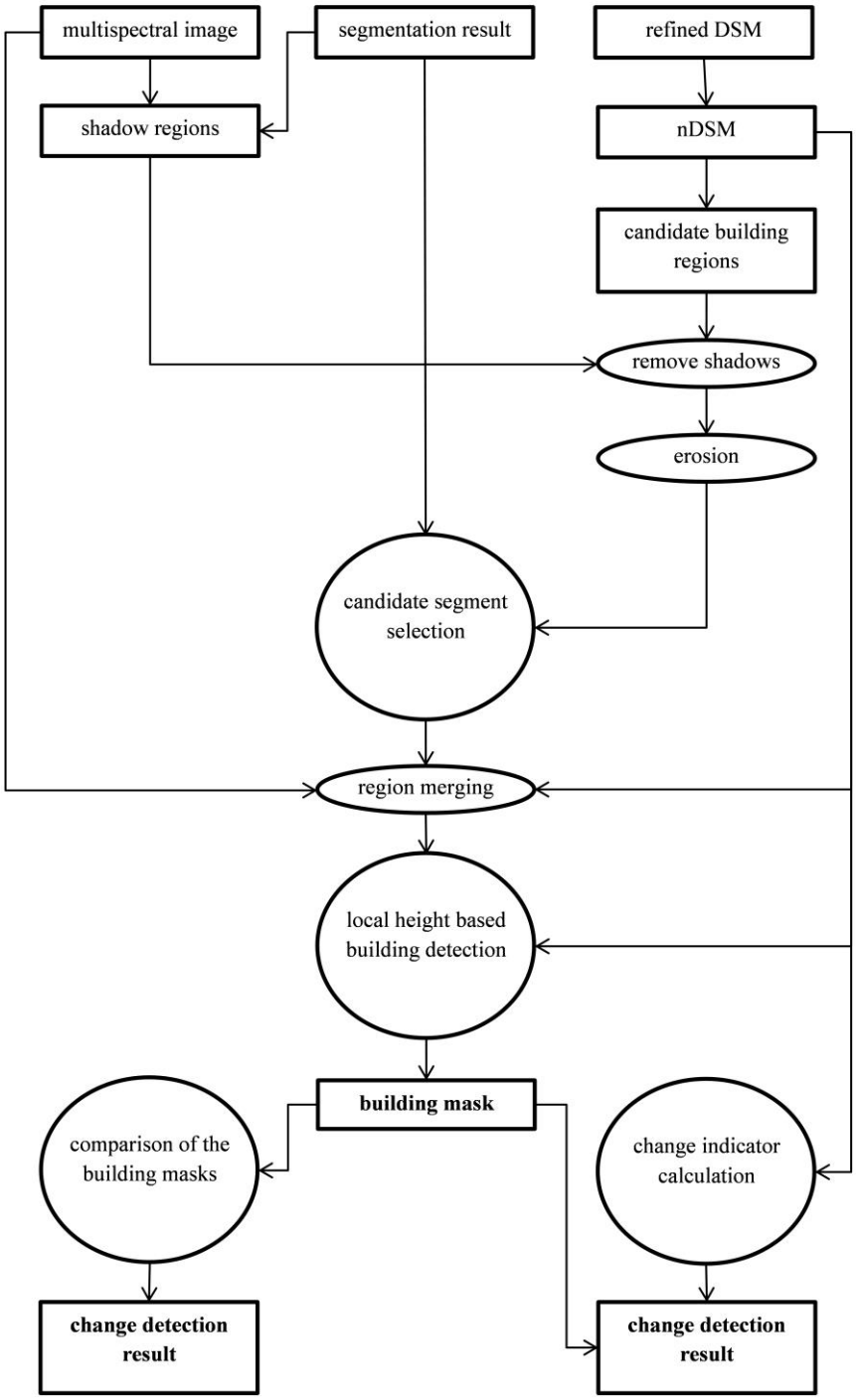

Figure 4-2: Workflow part 2 (building footprint extraction and change detection)

4.2 DSM refinement

As the quality of the DSM has a large impact on the results of further steps, in the first step, the DSM quality has to be analysed and improved

Therefore, there is at first segmentation needed which is based on intensity gradients that are gained from the panchromatic input image. This initial segmentation is further improved by region merging. Based on the obtained refined region map, DSMs are refined by taking the average height of selected regions.

4.2.1 Gradient based segmentation

As a first step for the gradient based segmentation, noise is reduced from the panchromatic image with a median filter (see Figure 4-3 (b)). Subsequently outliers that are farther apart from the average value than a certain amount of standard deviations are given a defined maximum value if they are too large, or a minimum value in case of being too small. The range of values is afterwards normalized with stretching, which is leading to a higher contrast within the given range. The benefits of this procedure called "*contrast stretching*" are shown in Figure 4-3 (c).

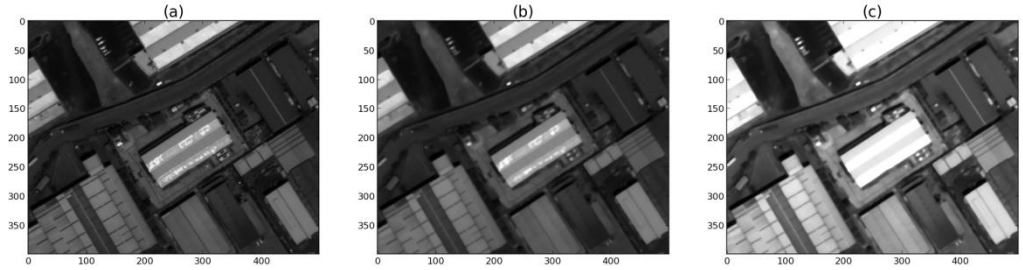

Figure 4-3: (a) Original panchromatic image. (b) Panchromatic image after median filtering. (c) Panchromatic image after contrast stretching.

It could clearly be seen, that after the median filtering building roofs and streets are more homogenous, which will lead to more uniform segments. In addition, the increased contrast is providing a better distinction between buildings and their surroundings.

After that, the Sobel gradients in horizontal (Sx, Figure 4-4 (b)) and vertical (Sy, (c)) direction are calculated and added up.

(4)

$$Sxy = Sx + Sy$$

As can be seen in Figure 4-4 (d), this sum (Sxy) is already highlighting the building boundaries quite well.

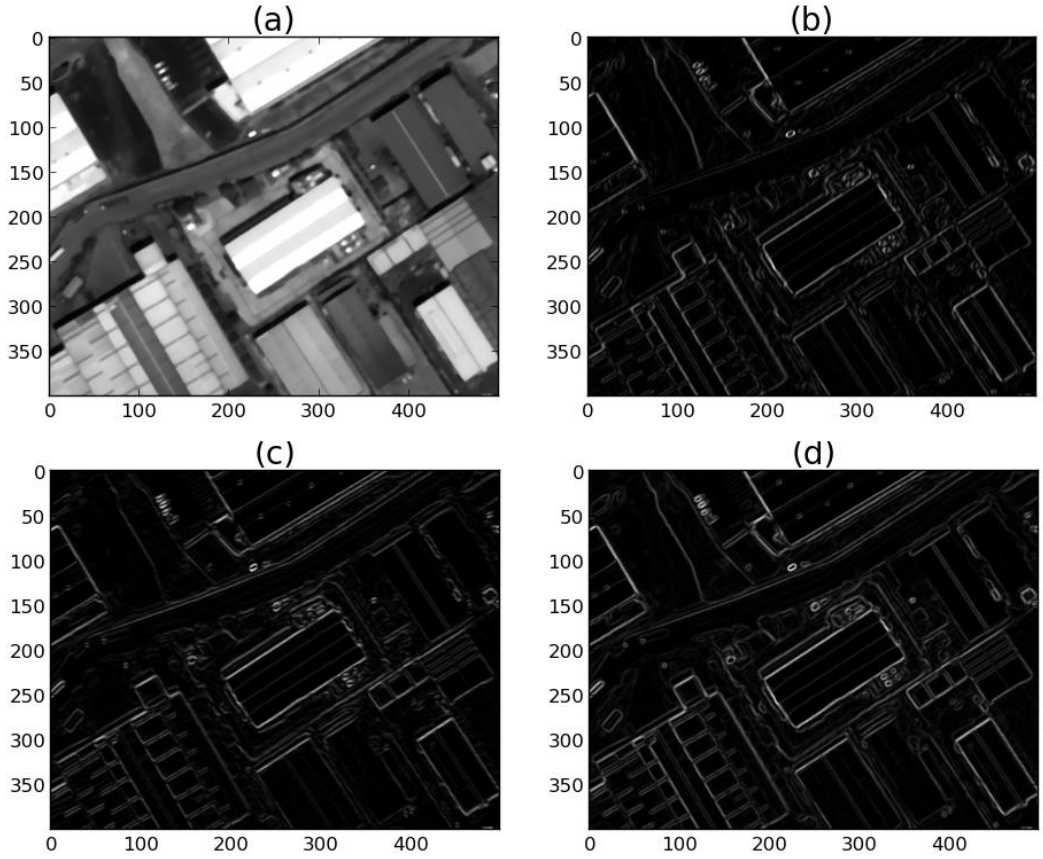

Figure 4-4: (a) Original panchromatic image. (b) Sobel gradient in horizontal direction. (c) Sobel gradient in vertical direction. (d) Gradient in horizontal + vertical direction

In order to extract the boundary pixels, thresholding is required to achieve the result as shown in Figure 4-5 (a). A threshold (T_{Sobel}) is set which divides between boundary pixels (bo) and area/background pixels (ar) . Thinner or thicker boundaries can be adjusted for individual purposes, however, trials showed that using the weighted average gradient $(\omega * \mu S_{xy})$ value over the whole image is delivering good results.

(5)

$$T_{Sobel} = \omega * \mu S_{xy}$$

For the next step, the labelling of the binary image, it is important that buildings are completely surrounded by boundary pixels (bo) (pixels with a gradient value above T_{Sobel}). Otherwise patches that are not belonging to each other will mistakenly be part of the same segment, as can be seen in Figure 4-5 (a, b) marked with red. In Figure 4-5 (b), every enclosed background pixel area (ar) becomes its unique label as represented with the different colours, while the boundary pixels (bo) are remaining the value zero (blue) in the resulting image.

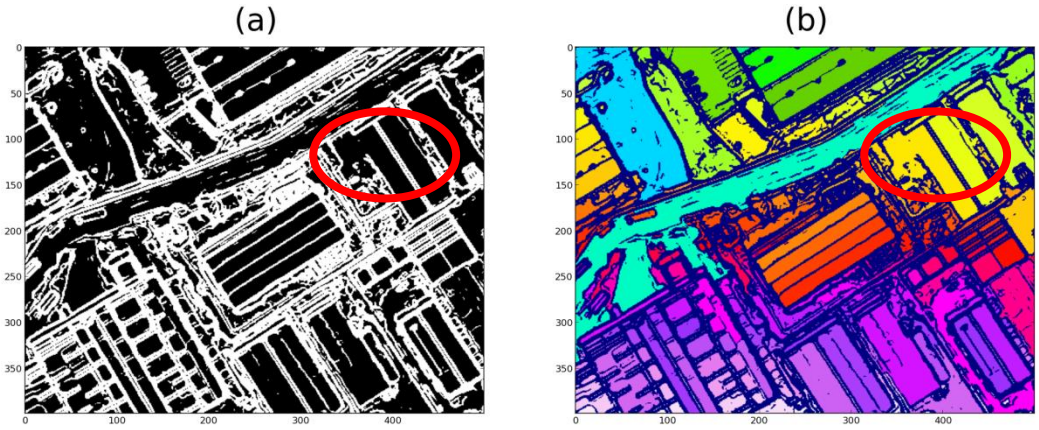

Figure 4-5: (a) Boundaries in a binary format. (b) Segmentation after labelling.

As the boundary segments (bo) are strongly connected with each other (see Figure 4-5 (a)) and so present as single objects, they need to be split up into smaller segments. Therefore the boundaries (bo) are labelled in a separate image as well as the background areas (ar) before. For that, a method is used that determines the orientation of the boundaries.

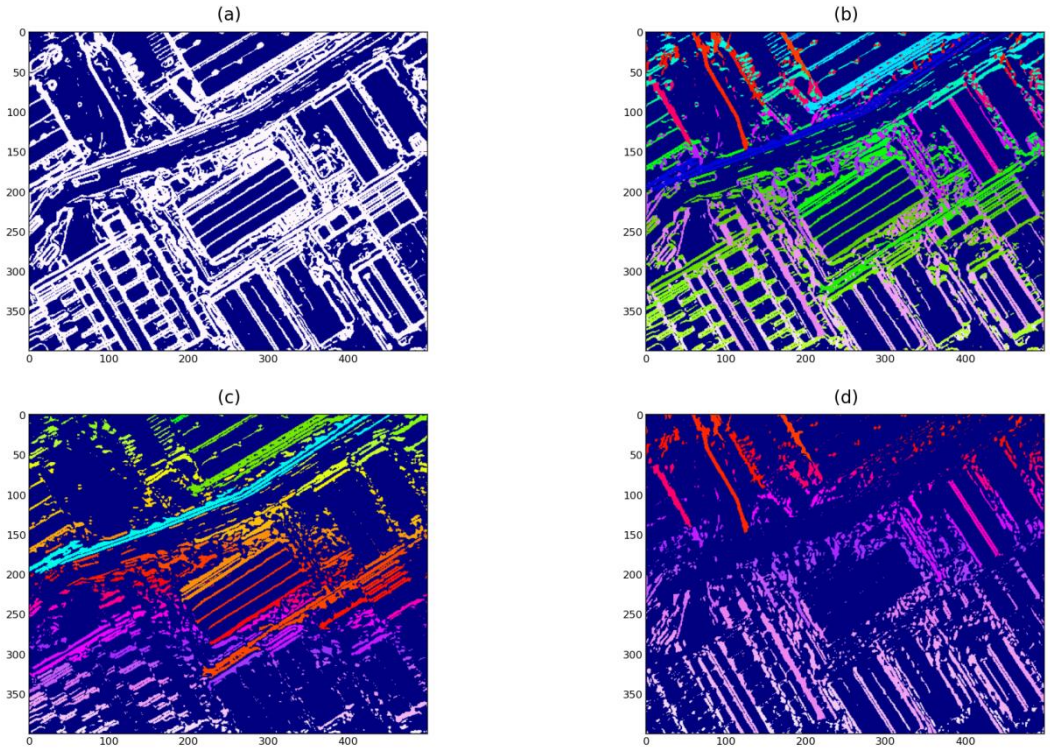

Figure 4-6: (a) Original binary boundaries mask. (b) Divided boundary segments. (c) Labelled boundary segments with a larger vertical component. (d) Labeled boundary segments with a larger horizontal component

In order to obtain the orientation of each boundary, the magnitude of the Sobel gradient in horizontal and the gradient in vertical direction are compared. This makes it possible to

distinguish between basically two cases. For having the gradient in accuracy with several fractional digits, the case of both being completely the same could be neglected and so, either the horizontal gradient is bigger (2), or the vertical component is (1).

(6)

$$1)\ Sy > Sx$$

$$2)\ Sy < Sx$$

This binary information is then transferred into two images, each for one of the cases. These are further labelled as is shown in Figure 4-6 (c) for a dominant horizontal component and (d) for a larger vertical gradient. After adding up both labelled images (c, d), all boundary segments are in one image, but split up, wherever the orientation of the boundary changes. This result is illustrated in Figure 4-6 (b), whereas (a) is showing the connected boundary segments from the beginning.

The labelled area segments (I_{area}) and the labelled boundaries ($I_{boundary}$) are combined afterwards to one image ($I_{labeled}$). This could be done with summation, since the background values in both images have the value zero. To avoid having labels twice, the labels of the boundaries are raised ($I_{boundary\_raised}$) by the amount of the maximum label of the first image ($Max(I_{patch})$). In addition, based on this method, the boundaries can be easily distinguished as they are labelled with values larger than ($Max(I_{patch})$.

(7)

$$I_{boundary\_raised} = \{I_{boundary} \mid I_{boundary} > 0\} + Max(I_{patch})$$

(8)

$$I_{labeled} = I_{boundary\_raised} + I_{area}$$

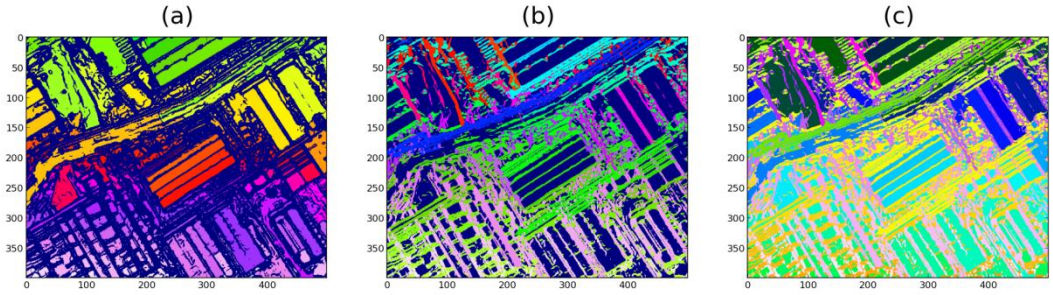

Figure 4-7: (a) Labelled area segments. (b) Labelled boundaries. (c) Combined segments.

4.2.2 Region merging

In this section, the segmentation result is further improved by merging. Several procedures are adopted for this purpose.

As shown in Figure 4-8 (b), many small segments are produced by the gradient based segmentation. They mainly locate at the boundaries of objects. The first task it to merge these small segments into one of their neighbours. Thereby, at first small segments with a size below a predefined threshold are assigned a “no data” – value, zero for instance (see Figure 4-8 (c), small blue areas).

These areas are labelled in the next step, which makes it possible to iterate them. By doing so it is checked if they are encapsulated by other segments (Figure 4-8 (d)). For that, the labels of the four direct neighbours of each pixel (as marked in Figure 4-9 with the red cross) in a segment are written into a list. Obviously, this is just possible with pixels that have four neighbours, which is not given at the boundaries of the image. This makes it necessary to extend the array size by the amount of one pixel (green pixels in Figure 4-9). The value of the new extension pixels thereby must be unique in the image, which makes it possible to remove them afterwards from the list with the neighbour pixels. Furthermore, also those entries are deleted, which do have the same label as the current examined segment. After that, there are just the surrounding values of the segment left.

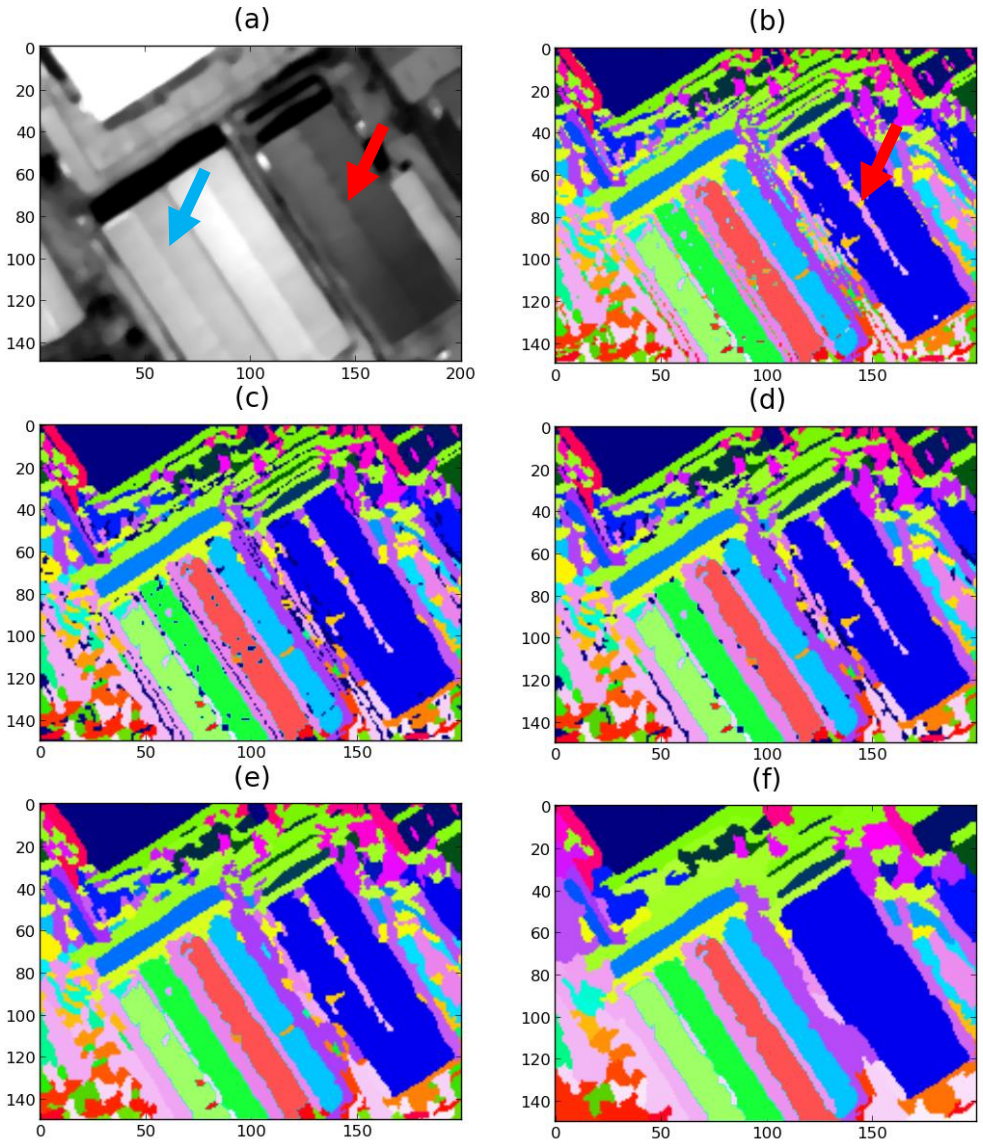

Figure 4-8: (a) Original panchromatic image. (b) Initial segmentation results. (c) After excluding small regions. (d) After merging the encapsulated small regions. (e) After assigning pixels to most similar neighbour. (f) After merging based on encapsulation

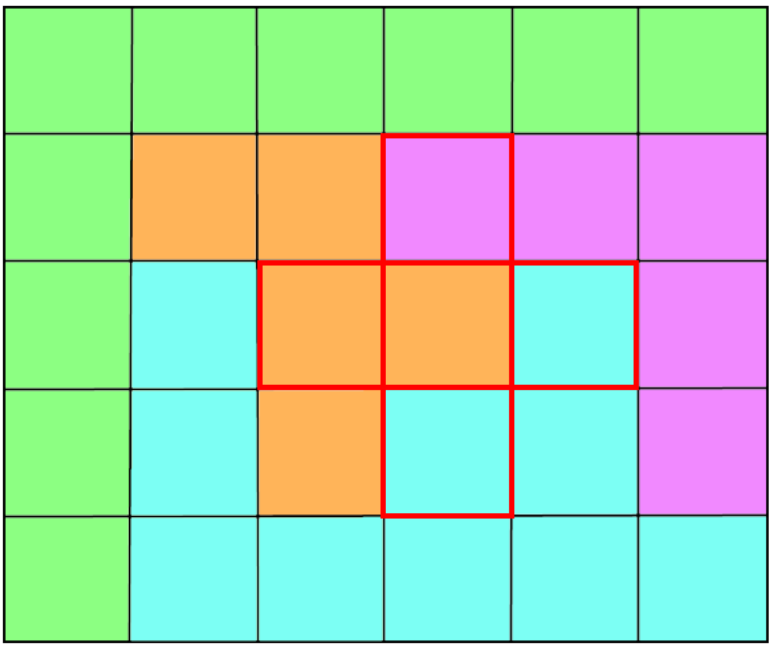

Figure 4-9: Merging based on encapsulation. Green pixels are representing the extended boundary pixels. The Orange pixels are merged to the blue area.

The quantity of each neighbour label in the list is counted subsequently. By dividing the count of the biggest neighbour with the total number of list entries, a ratio is calculated that represents the encapsulation by another region.

For example, the orange area (O) in Figure 4-9 should be examined. After each label in the defined Neighbourhood from each pixel in O is written to a list, there are eight orange, three green, two purple and 7 blue pixels. Excluding the orange and the green ones, leads to nine pixels in total. Consequently O is enclosed though 7/9 by blue and by 2/9 by purple.

If the highest ratio is larger than a threshold ($T_{encapsulaton}$), the encapsulated region is assigned the label of the most surrounding segment. Otherwise, no merging is performed. This leads to a better representation of buildings in the region map, as those building areas are now more homogeneous than before (shown in Figure 4-8(d)).

However, there are still some “nodata” – areas left (small blue areas (d)). These are filled pixel wise with the value of the most common label in a three by three window. If there are just “nodata” pixels within this window, than the size of it is raised until there is pixel with a label. The result of this is shown in Figure 4-8 (e).

As is shown in (e), the number of segments is largely reduced after the region merging step. However there are still small segments surrounded by larger ones. To further improve this result, the third merging step is proposed. In this step all segments are checked instead of only small size segments. This leads to more homogenous segments as is shown in Figure 4-8 (f). A good example for the improvement is the blue building,

which is marked with a red arrow. After this procedure it is well represented by one single segment, as it is the goal for building footprint extraction.

It has to be noted, as a more common case, that buildings are divided into several main parts, which could also be seen in Figure 4-8(a). The representative building is thereby marked with a blue arrow. This separation is due to ridge lines on the pitched roofs. These are in the segmentation represented as boundary areas, for having a large gradient. This could be formalized to the case subscription that two area segments are connected by a boundary segment, as is shown in Figure 4-10. If they are belonging to the same building, all three parts should be merged together to one segment.

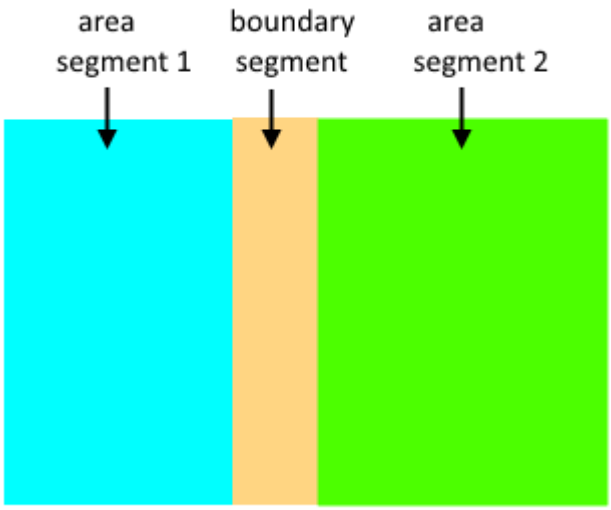

Figure 4-10: Initial situation for triple merging

For that, all boundary segments (purplish coloured in the referred building area in Figure 4-8) are checked if they have two main area neighbours. This is done by the same neighbour detection method as in the merging procedure before. In case of the first requirement being fulfilled, it is subsequently figured out if all three segments do have a height average (see Figure 4-11 (a)) difference smaller than a set threshold (T_{height}).

(9)

$$T_{height} > Max\{D_{height\ 1,2}, D_{height\ 1,3}, D_{height\ 2,3}\}$$

Furthermore the average intensity values of the two area segments are checked before merging (Figure 4-11 (b)) value, in order to be merged. The boundary segment is not used in this step, as the boundary regions have always a different intensity per definition in order to have a high gradient.

(10)

$$T_{intensity} > D_{intesity\ 1,2}$$

As the panchromatic image was normalized before, it is satisfying to set the threshold ($T_{intensity}$) to a fraction of the total range. If all affordances (see Equation (9), (10)) are fulfilled, then the three segments can be merged to one. The result of this procedure is shown in Figure 4-11 (c).

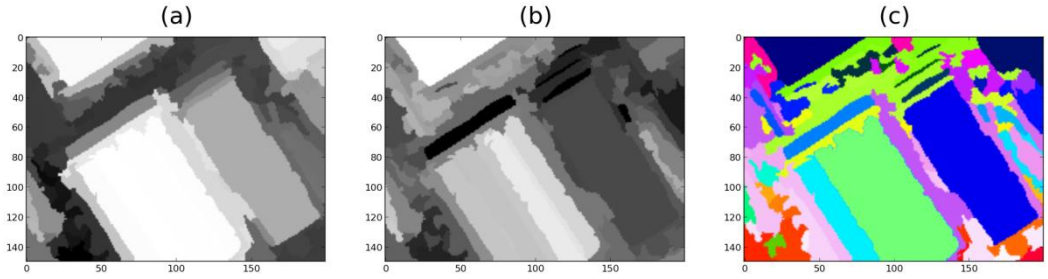

Figure 4-11: (a) Average height over the Segments. (b) Average Intensity over the Segments. (c) Resulting segmentation.

As can be seen in Figure 4-11, the uniformity and representation of buildings in the segmentation map was significantly improved after the merging procedure. Based on this prepared segmentation map, the quality of the DSM can be improved.

4.2.3 Creating the refined DSM

One of the remaining problems after the merging is under segmentation, as shown in Figure 4-5 and Figure 4-12 (c). This is caused mainly by the similar textures between building roofs and streets. As shown in (Figure 4-12 (b)) the building roof and ground are incorrectly recognised as one segment (c). Therefore, the prepared region map cannot be directly used for the DSM refinement procedure, and it is necessary to check if the resulting height changes are reasonable.

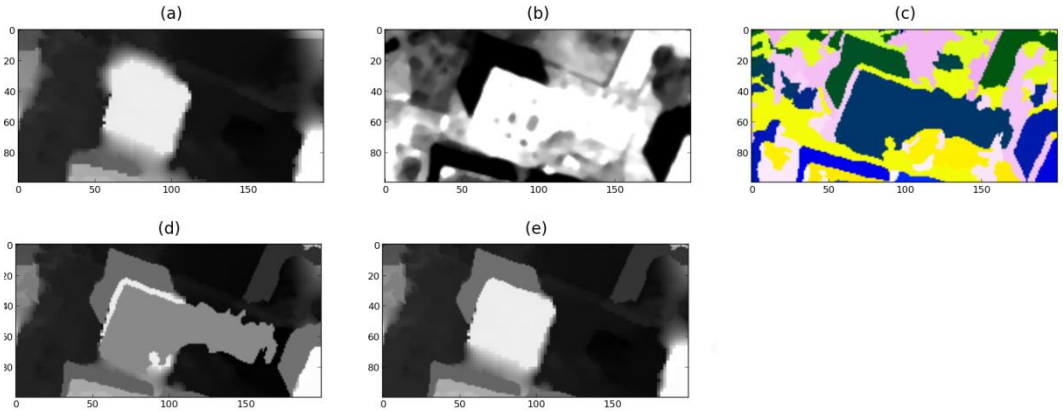

Figure 4-12: (a) Original DSM. (b) Contrast stretched panchromatic image. (c) Segmentation result. (d) Initial refined DSM result. (e) Refined DSM result with error check

Therefore, in this step the average height over the segments are compared to the original height. If the mean standard error between this average height ($H_{average}$) and the original height ($H_{original}$) is above a threshold T_{change}, then the segment is not used for the refined DSM. Concluding the original height values are kept in the refined DSM, as is shown in Figure 4-12 (e).

(11)

$$T_{change} > MSE(H_{original}, H_{average})$$

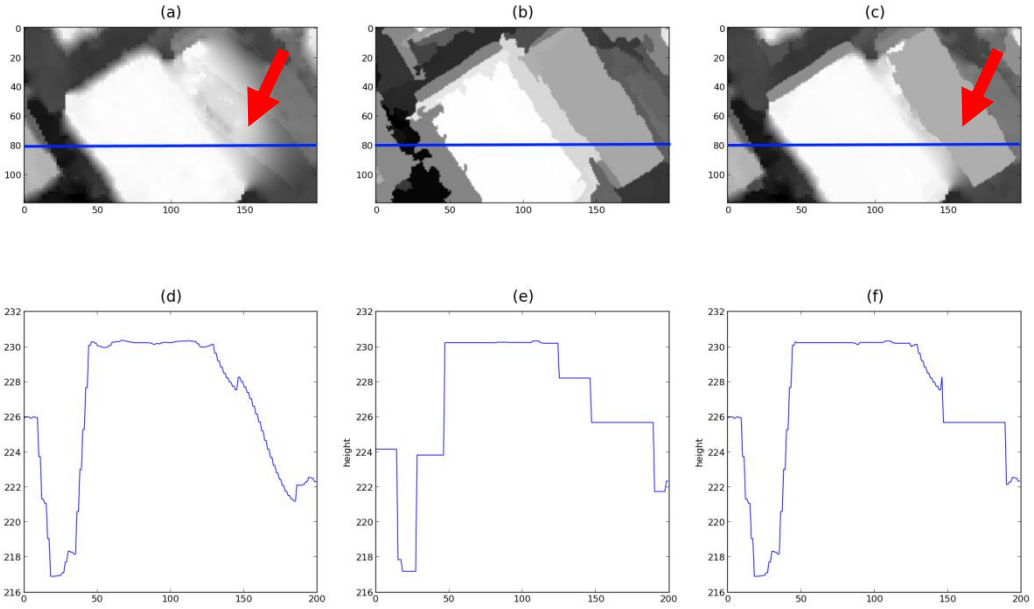

Figure 4-13: (a) Original DSM. (b) Refined DSM with average over boundaries. (c) Refined DSM result. (d,e,f) Height profiles of a,b,c at the blue line.

In general, the refinement is based on averaging the height over reasonable area segments and overwriting the referring areas in the original image, which is shown in Figure 4-13 . Boundary segments are therefore not used for the refinement, as they are often connecting different areas, as could be seen in Figure 4-8 and Figure 4-11 (c), the result would concluding be more vulnerable to mistakes and less precise as shown in Figure 4-13 (b). Indeed, it would be possible to correct this with the error check procedure, if the threshold for wrong change features would be low enough. However, this again could eliminate positive DSM refinements, as are shown in the height profiles in Figure 4-13 (d-f) where was marked with the red arrow in Figure 4-13 (c). The referring panchromatic image is thereby shown in Figure 4-8 (a) also marked with a red arrow.

Summarizing, it could be stated that the focus of this method is to add as much correct information to the DSM as possible, by risking as few as feasible in terms of errors.

4.3 Segmentation

After the DSM is refined by using the panchromatic images it can be further contributed in a more advanced segmentation procedure. In this section, the multispectral and panchromatic images and the refined DSM are used as features for segmentation.

4.3.1 Data preparation

As is shown in Figure 4-1, the refined DSM is one of the input images. Furthermore the panchromatic and a multispectral image, that should help to distinguish between buildings and their background, are used as well. These data needs to be prepared for the later steps at first.

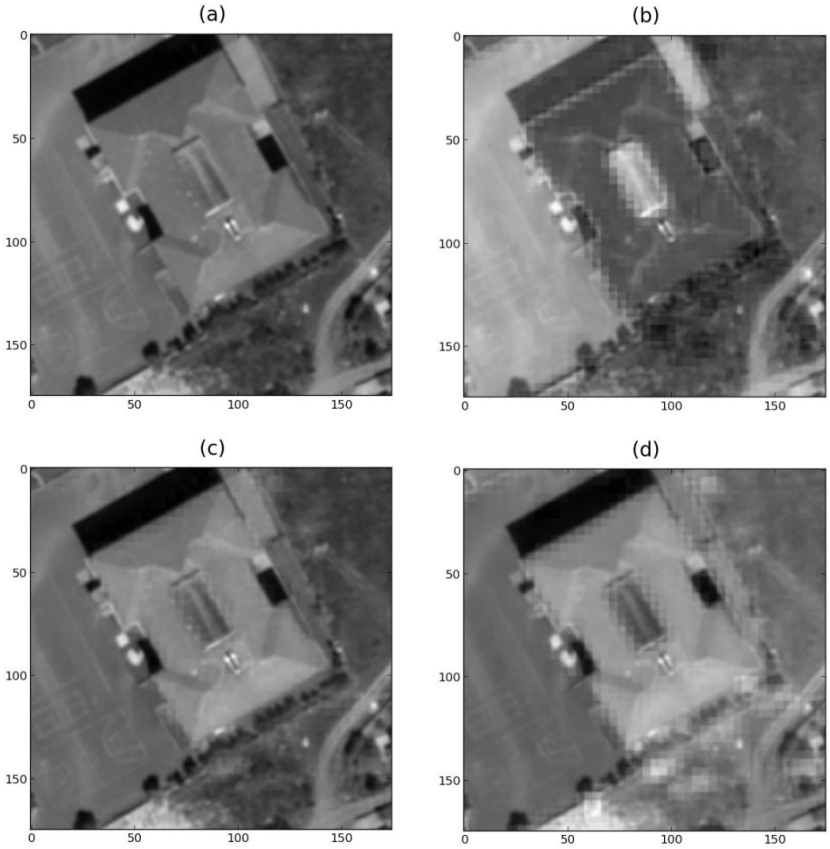

Figure 4-14: (a) Original Panchromatic image. (b - d) Multispectral channels

In terms of that, the quality of the multispectral image is improved with pansharpening. Then the principal component analysis (PCA), is performed on the pansharpened image (Figure 4-14 (b-d) to extract the main information (principal components) from the channels. It leads to a compression of information, whereby the number of channels is reduced (Adbi and Williams 2010). Additionally, it leads to a faster calculation of the mean shift segmentation, which is strongly influenced by the number of input features. Furthermore it makes the procedure more robust, because by using the PCA, information

from all channels is used and there is no need for selecting and weighting a few channels from it. Based on PCA, the redundant information (see Figure 4-14 (b-d)) is eliminated, while the variant features which are improving the segmentation are kept.

The result of the PCA is consisting of several components, with a decreasing amount of information. Therefore it is sufficient to use just the first three channels for the succeeding steps.

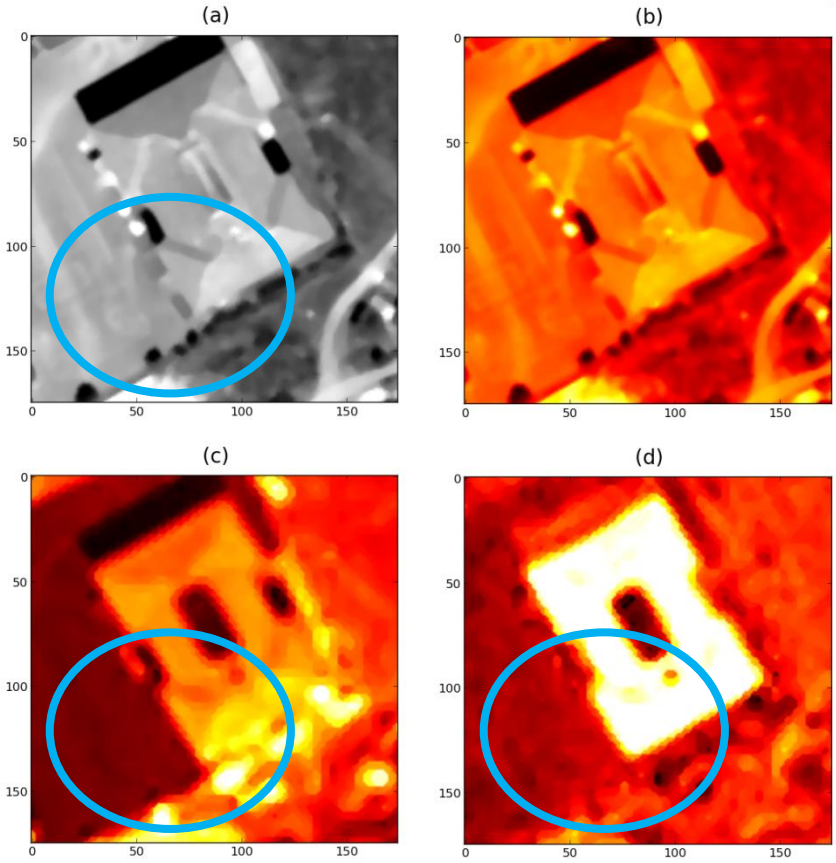

Figure 4-15: (a) Contrast stretched image. (b - d) Contrast stretched PCA – channels.

After contrast stretching with the in 4.2.1 explained method, the differences between the buildings and their background are largely increased. By comparing this to the contrast stretched panchromatic image, as shown in Figure 4-15, the gain of information due to the multispectral image is salient. While the panchromatic image (a) has just a minor intensity difference between the background and the building (marked with a blue circle), especially the second (c) and third (d) PCA – channels are providing a good foundation for the distinction in this case. In other scenes, the opposite, that the panchromatic image is better for the differentiation is also a common pattern. Therefore, it makes it possible to achieve reliable segmentation results by using all the information together.

4.3.2 Mean shift segmentation

The mean shift segmentation is one of the key parts of the whole procedure and is able to use multiple input features, to find clusters in the data (Comanicu and Meer, 2002).

For this, it is at fist needed to convert the input images to the feature space. Each pixel thereby is represented as one data tuple. These are having several features, whereas their number is given by the number of input images. If the spatial location of the former pixel is used, then x and y position are features of the tuple as well. Figure 4-16 thereby shows the scene from Figure 4-15 in feature space, although just a few of the features of each data tuple could be depicted.

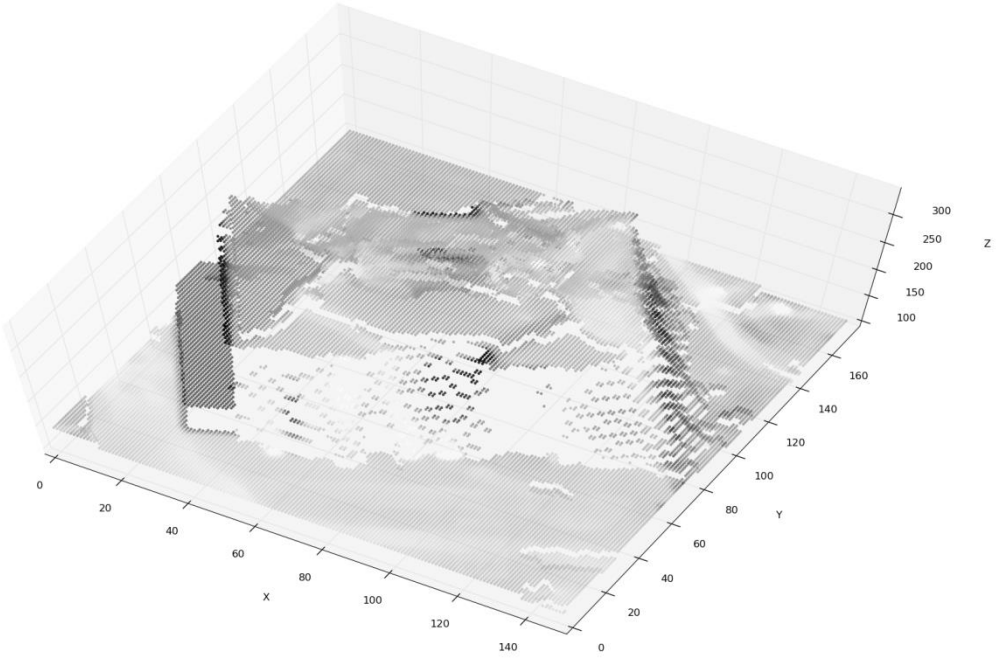

Figure 4-16: Image in Feature space. (X, Y, Z - values of tuples + intensity)

It is furthermore visible in Figure 4-16, that there are clusters in the feature space. These could be detected with the mean shift algorithm, as is shown in Figure 4-17.

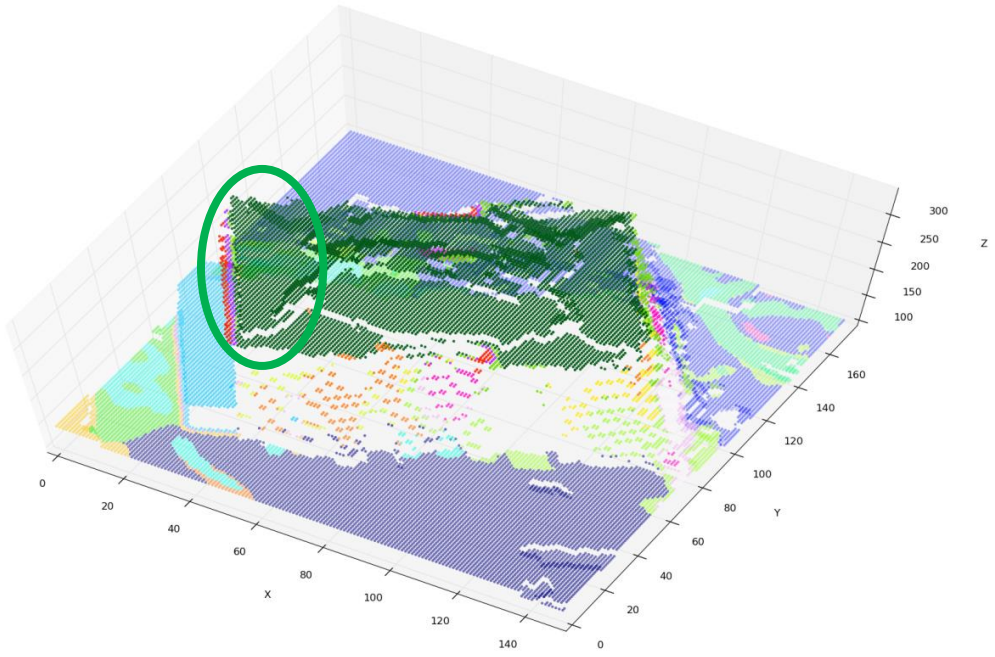

Figure 4-17: Clusters in the feature space gained with mean shift segmentation

The size of the determined clusters is thereby dependent on the shape and size of the used kernel and the range of the feature space. The used kernel is uniform, which means that it has the same extent in all dimensions. If it is wanted to put more weight one some of the features, this could be done by adjusting their scale range. Raising the extent of a feature thus leads to a higher weight of it, while a smaller range leads to a minor impact. This coherence is shown in Figure 4-18 , where the range of the height is decreased. A good example for this area marked with a green ellipse in Figure 4-17 and Figure 4-18. While the larger impact of the height in the first image was leading to a more homogenous roof of the building, a smaller influence is causing a relatively higher impact of the other input features. This leads in this case to more fragments.

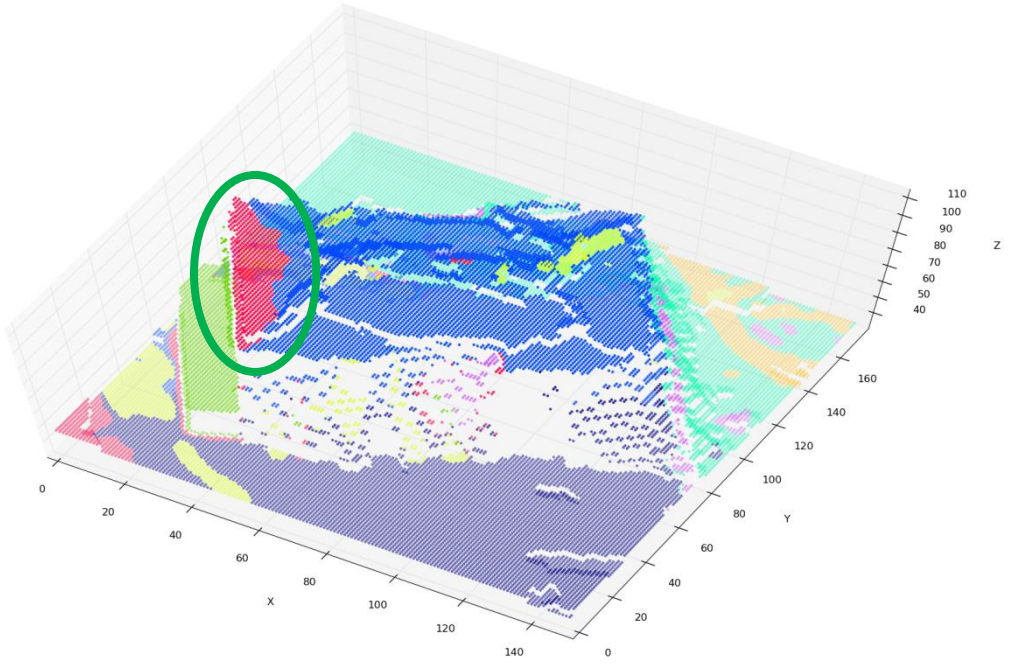

Figure 4-18: Segmented Feature space after decreasing the scale range of the height

In general, a higher scale range is leading to smaller size - and more numbers of clusters, for the differences within the data are getting bigger. The same effect occurs by expanding the kernel size. Here the smaller kernel is more sensitive to discontinuities in the data.

After the Clusters are determined, the data needs to be transformed back to image. Thereby, just the cluster information from each tuple and later pixel is relevant. The resulting images from both previous described clustering approaches are shown in Figure 4-19 .

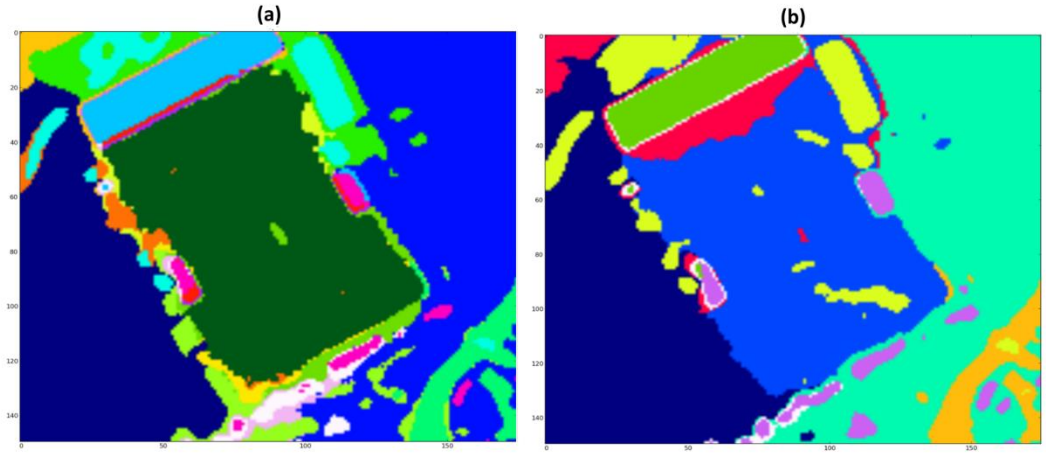

Figure 4-19: (a)Segmentation based on Figure 4-17. (b) Segmentation based on

4.3.3 Segmentation refinement

For the following building footprint extraction it is necessary to refine the segmentation at first, for too much noise and too many small segments, as can be seen in Figure 4-19 (a) would have a negative influence on the later results. Despite that, there is no extensive region merging for all segments in the image performed at this step, for merging could be performed more efficient later, when the building candidate regions are iterated. Then just the possible building regions are examined and not the whole image with a lot of parts that are not of interest.

In this step, all segments that are smaller than a certain threshold T_{size} are merged to their best fitting neighbour k_{Hmin}. This is by determining the size of the segments S_{size} at first. Then the neighbour regions $\mathrm{k}_i\ (i = 1,\ldots,n\ , n = number\ of\ neighborhood\ regions)$ from region S are searched. Finally, they are merged to the neighbour with smallest height difference H_{sk_i}.

(12)

$$H_{s,k_i} = \left|\mu H_s - \mu H_{k_i}\right|, \qquad S_{size} < T_{size}$$

(13)

$$S_{label} = \mathrm{argmin}(H_{sk_i})$$

Furthermore, encapsulated areas are merged, for they are assumed to be part of the segment that surrounds them. The Method used is the same as explained in 4.2.2. Thereby Figure 4-20 shows the significate reduction of noise in the segmentation result after the two steps.

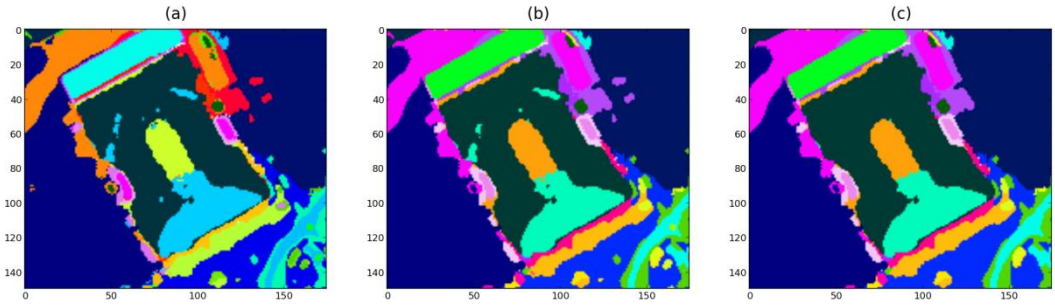

Figure 4-20: (a) Segmentation result. (b) Segmentation after merging small areas. (c) Segmentation after merging based on encapsulation.

4.4 Building footprint extraction

The workflow of the following method is shown in Figure 4-2. As can be seen, it is divided into two main parts. At first the initial building candidate areas are determined, after that building oriented merging is performed.

4.4.1 Initial building candidate selection

This step tries to find possible building locations with the help of the refined DSM. Therefore, it was normalized with a morphological reconstruction function provided by the DLR. In the resulting normalized DSM (nDSM), which is shown in Figure 4-21 (c), all pixels in ground region have a value of less than about two meters, while objects like buildings have the relative height values in reference to the ground. The difference in the range between (b) and (c) thereby could be seen best in the scalebars, marked with a red arrow.

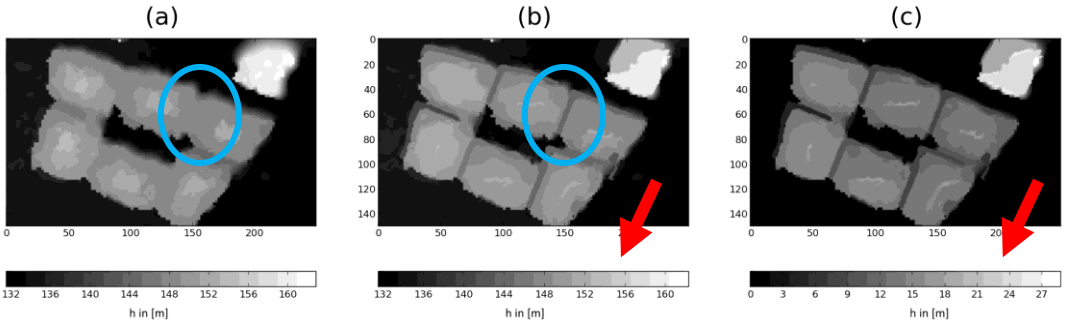

Figure 4-21: (a) Original DSM. (b) Refined DSM. (c) Normalized refined DSM

Subsequently, it is assumed that buildings are having a certain minimum height. For that, all pixels with a height value smaller than a given threshold could be eliminated, as is shown in Figure 4-22 (a, b), marked with a blue circle. All non-zero values are then transferred into a binary array and labelled (c). Each segment with a size smaller than a threshold is again assigned to the background, while the other pixels are again transferred to a binary array (d), which is needed for further erosion.

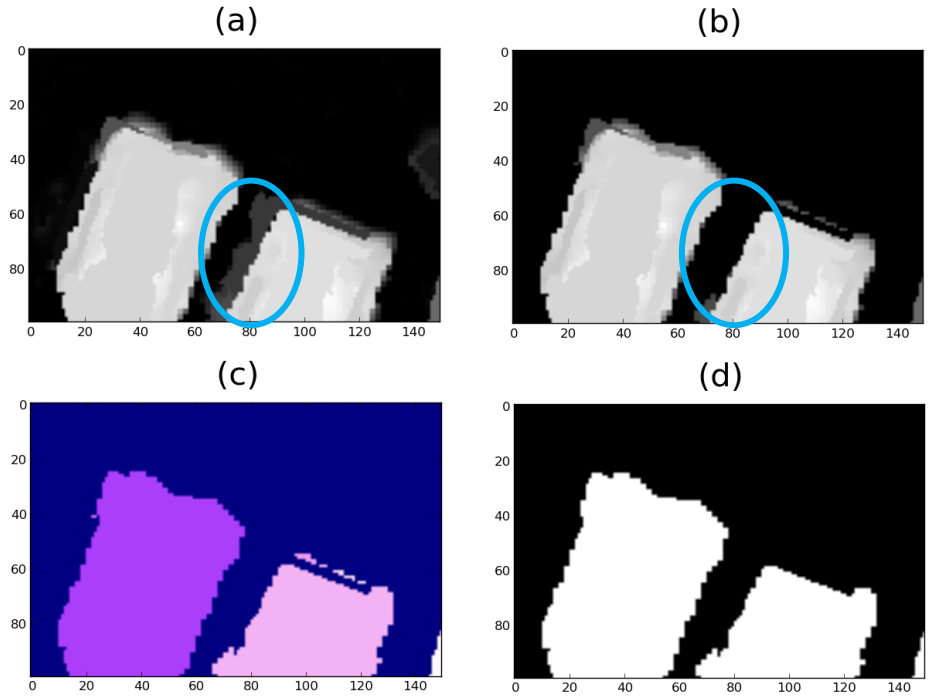

Figure 4-22: (a) nDSM. (b) After assigning the height based threshold. (c) Labelled building candidate area. (d) Binary building candidate area.

However, it is quite difficult to give a proper threshold value in this step. A higher threshold value leads to possible miss detections. While with a lower threshold, some buildings may be wrongly connected together limited to the DSM qualities. Therefore, shadows are adopted in this step to separate these buildings from another. The shadows were thereby detected with an already implemented function from the DLR. The calculation is thereby based on multispectral data. Despite there might be some buildings with shadows on their roofs, the assumption was made that shadows are not part of buildings. This helps in many cases to separate buildings that are standing close together, as shown in Figure 4-23 (a).

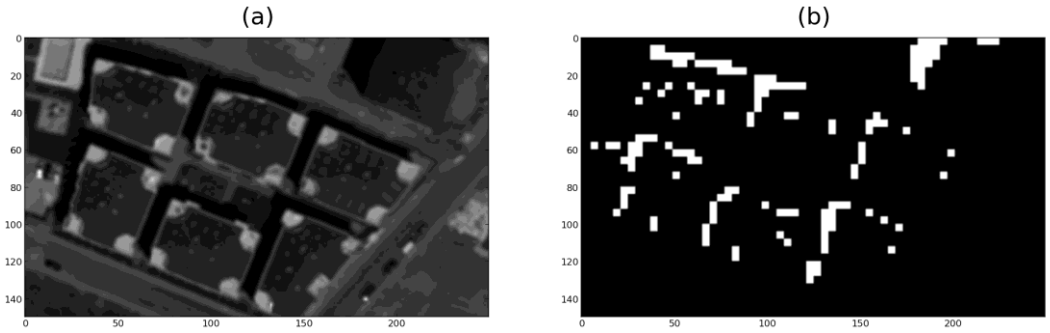

Figure 4-23: (a) Panchromatic image. (b) Calculated shadows.

Usually the accuracy of the DSM in shadow regions is not good (see Figure 4-21 (a), blue circle), for completely black areas could not be matched properly in the DSM generation

procedure. The often shadowed space between buildings, for that is often blurred and so there is no clear delineation of buildings in the DSM recognizable. Due to the coherence between the shadow regions and blurred areas, eliminating the shadows from the building candidate mask helps to separate buildings.

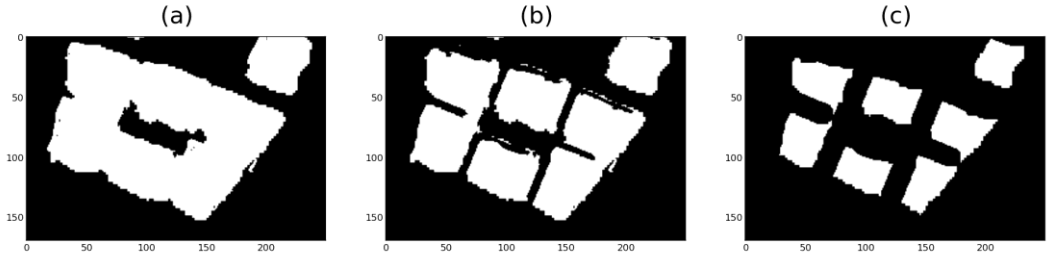

Figure 4-24: (a) Initial building candidate areas. (b) Building candidate areas after exclusion of shadow regions. (c) Building candidate areas after erosion.

The location of the shadows (Figure 4-23 (b)) was thereby calculated from the multispectral image by examining the spectral characteristics of all single pixels. After that, the shadow cover rate is determined. If the shadow cover rate is above a certain threshold, the whole area is classified as a shadow. Therefore, these areas are eliminated from the candidate mask, as is shown in Figure 4-24 (b).

As next, it is further tried to separate mistakenly connected areas from another with morphological opening. For this, morphological erosion is used first. Thereby the areas are eroded from the outside. This helps especially when there are small connection corridors between some main areas, as is shown in Figure 4-24 (b). The areas are shrinking and ideally splitting up into isolated parts that are representing the single buildings (see Figure 4-24 (c)).

4.4.2 Building oriented region merging

The generated building candidate mask is then labelled for the further process.

At first, the areas are brought back close to their original size with dilation by the same amount as the erosion before. This is shown in Figure 4-25 (b),

Next, all segments of the segmentation map from 4.3, are selected that overlay with the current building candidate area. Segments, which have a minimum predefined coverage of the area are selected, and named as building candidate segments. While segments with relatively less cover rate on the building candidates regions will not be selected. Figure 4-25 (c) thereby shows the building candidate segments. Without the segmentation refinement steps in 4.3.3 , some of the small segments at the boundaries with a too low coverage might have been lost, due to the small accuracy of the building candidate areas.

These selected segments are then tried to merge. At first, all the segments with a size below a defined threshold are merged to their best fitting neighbours in terms of height and intensity, as is explained in 4.3.3.

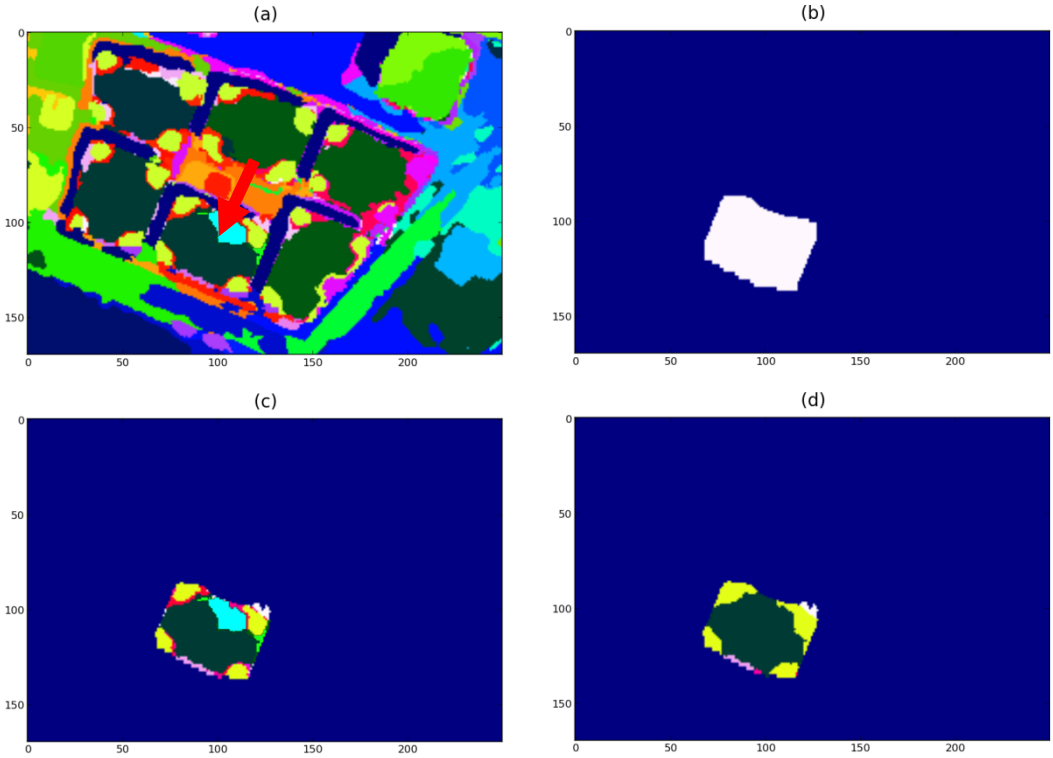

Figure 4-25: (a) Segmentation result. (b) Candidate area after dilation. (c) Segments within the candidate area. (d) Segments after merging.

After the small size segments are merged to larger ones, a more computational expensive method is used for building footprint extraction. In which the contrast increased PCA components from 4.3.1 and panchromatic image are used as well as the refined DSM. For that, the thresholds T_n for the similarity are needed to be set at first.

(14)

$$T_1 = x_1 * \sigma(I_1)$$
$$\vdots$$
$$T_n = \ x_n * \sigma(I_n)$$

This is as always done in dependency on the standard deviation of the given input image I_n, while x_n could be used to adjust the threshold. After that, all segments G_s are compared to their neighbours G_{k_i}. Thereby, the difference $D_{s,k,n}$ between the average values $\mu\mathrm{G}$ of the segments s, k, for the images n is calculated. The results are saved in the matrix Ω, whereby the columns are the input images $(1..\mathrm{n})$ and the rows are representing the neighbours of the segment $(1..\mathrm{k})$.

(15)

$$D_{s,k,n} = |\mu G_{s,n} - \mu G_{k,n}|$$

(16)

$$\Omega = \begin{bmatrix} \omega_1 * O_1^1 & \cdots & \omega_n * O_n^1 \\ \vdots & \ddots & \vdots \\ \omega_1 * O_1^k & \cdots & \omega_n * O_n^k \end{bmatrix}$$

Thus the average difference $D_{s,k,n}$ is compared with the referring threshold T_n. If the difference is smaller than the Threshold, then O_n^k is one, otherwise it is zero.

After that, the weighted sum W_k for each neighbour/vector (k) is calculated and the neighbour with the highest similarity W_{kmax} is selected.

(17)

$$W_k = \sum_1^n \omega_n O_n^k$$

(18)

$$W_{kmax} = \mathrm{arcmax}(W_j), j = 1, .., k$$

If W_{kmax} is larger than the before defined threshold $T_{weightsum}$, then the segments are subsequently merged to the referring neighbour k_{wmax} , otherwise no merging is performed. The merging results are shown in Figure 4-25 (d).

4.4.3 Building detection

After the merging procedure, building detection is performed. In order to get the final building footprint map. A local threshold method is proposed, which means instead of using a uniform threshold value for the whole test region, a minimum building height $T_{h,j}$ is automatically selected for each building candidate region $(1 \,..\, \mathrm{j})$ separately. For that, the average height μH_j and the standard deviation $\sigma(H_j)$ are calculated from the NDSM in those areas that overlay the segments of the candidate region. ω is thereby a weight to adjust the threshold. Based on this, the local threshold is defined that distinguishes between buildings and background.

(19)

$$T_{h,j} = \mu H_j - \omega * \sigma(H_j)$$

Every segment in the current candidate mask j is assigned to be a building if its average height value μH_s is higher than this local threshold.

(20)

$$\mu H_s > T_{h,j}$$

This provides benefits especially when the previous separation procedures failed, and there is more than one building within one building candidate region left, as is shown in Figure 4-26. In such a case it is possible to separate the large regions into single building areas (d) with this method. Therefore it is needed, that the areas between the buildings are having a lower average height than the buildings themselves (see (c)).

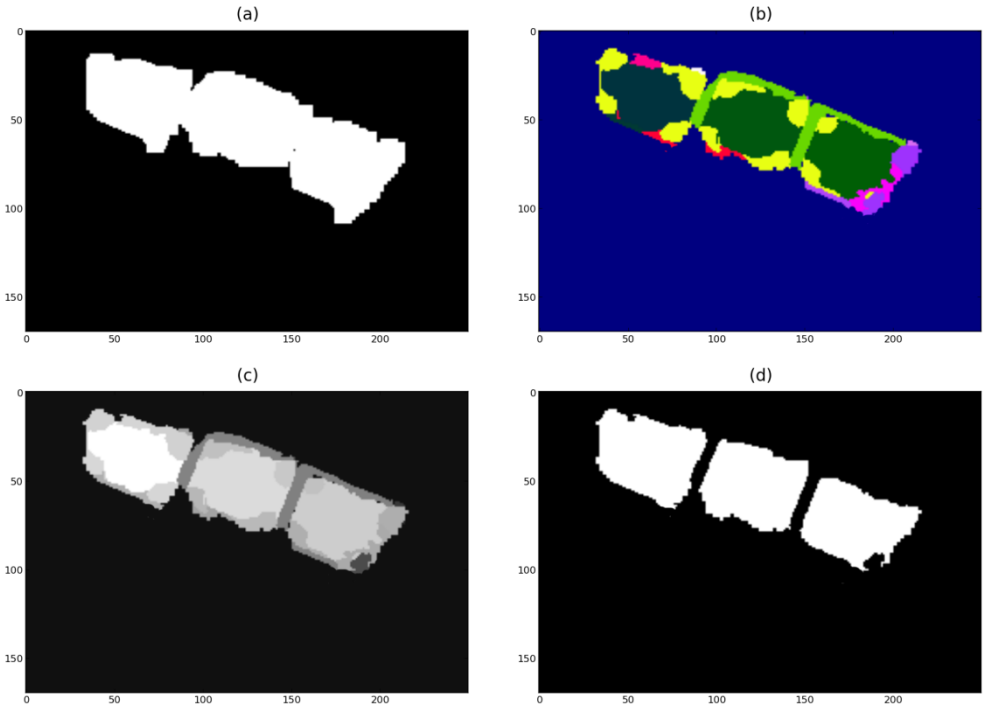

Figure 4-26: (a) Building candidate area. (b) Segments after merging. (c) Average height of the Segments. (d) Building footprint after applying the local threshold.

As this method relies on the nDSM, the quality of the nDSM has a large impact on the resulting building footprint. Therefore, the use of the refined DSM is leading to better distinction between buildings. Figure 4-21 (a, b marked with blue circle) thereby demonstrates exemplarily the better discriminability of single buildings in the same scene as in Figure 4-26. Moreover, the DSM refinement leads to a better segmentation in the previous steps, which has also a big influence on the resulting building footprint.

The detected building footprints might have some sticking out parts at the boundaries, which can be further sharpened by using morphological opening. This is done by using at first erosion and after that dilation using the same window size. After that, regions with small sizes are removed, boundaries flattened and unwanted leftover connections

removed. The gained building footprint can be subsequently used for the change detection.

4.5 Change detection

After the building footprint is calculated, region based change detection could be performed in two ways, as is shown in the workflow in Figure 4-2.

4.5.1 Post-classification building change detection

After the building footprints are extracted from images of two dates, building changes can be extracted by comparing them on the object level. Each building footprint in the later scene that is not overlaying one of the earlier scene, is supposed to be a new object. Therefore a minimum coverage needs to be defined. All buildings with coverage below that threshold are assumed to be new features. Both input images (a, d), building footprints (b, e) and the results (c, f) are shown in Figure 4-27.

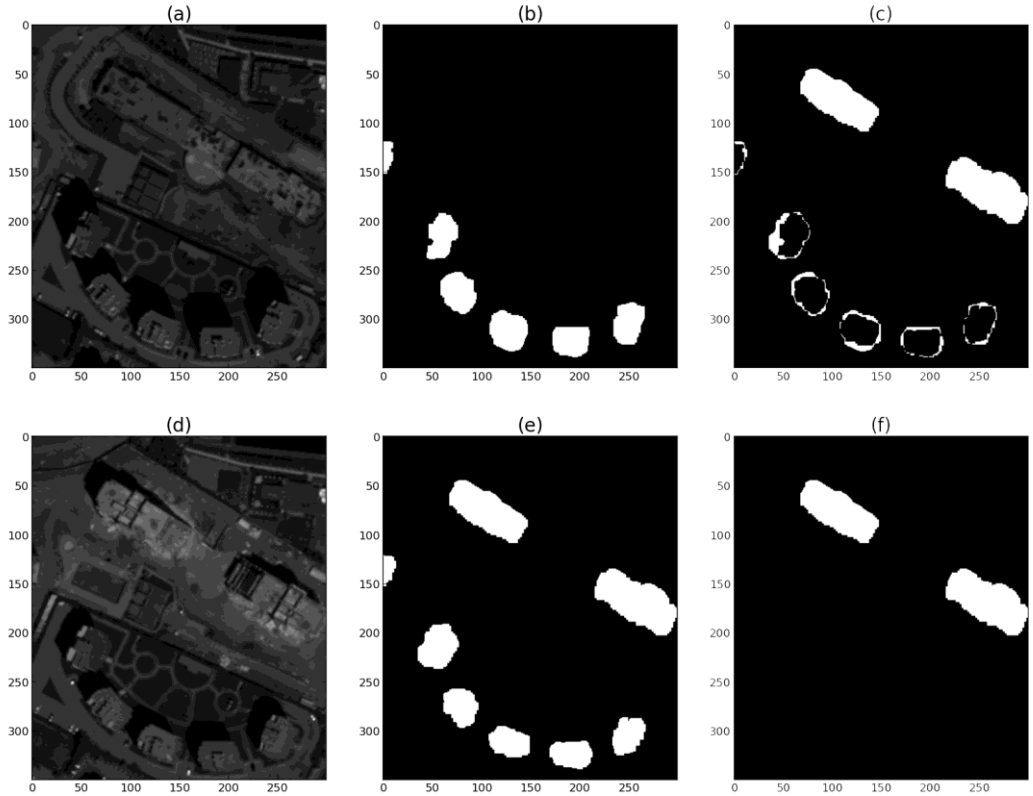

Figure 4-27: (a) Panchromatic image of date one. (b) Building footprint of date one. (c) Difference between the building footprints. (d) Panchromatic image of date two. (e) Building footprint of date two. (f) Change detection result.

4.5.2 Object building change detection

The other possibility to detect newly built buildings is to use just the building footprint of the second date. Additional a pixel based change indicator (Figure 4-28 (b)) needs to be

calculated from images of both dates. For that, an already existing function was used which was proposed by Tian et al. (2015). Thereby, the building change probability of each pixel is determined by using *believe functions* based on the DSM and the spectral images.

As next, a binary image change map is calculated out of the change indicators, whereby just those indicator values above a given threshold are transferred. After that, the coverage of the segments from the building footprint with the binary change map is calculated. Segments with a change indicator overlay above a before defined threshold are assigned as new objects. The result is shown in Figure 4-28 (c).

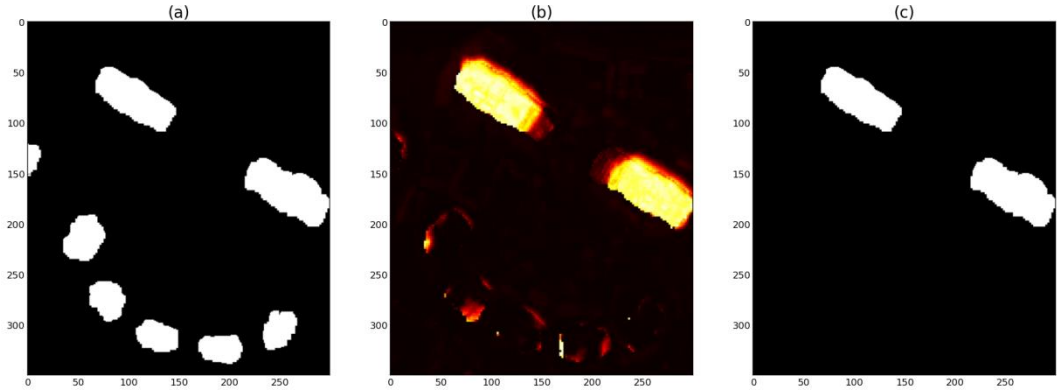

Figure 4-28: (a) Building footprint of date two. (b) Change Indicator. (c) Change detection result.

5 Experiment

Subsequently the experimental data are described, and results of each step are illustrated and evaluated.

5.1 Area and data

Two test sites (a, b) were chosen for the experiment. Thereby both are located in a mixed industrial and residential area in the east of Istanbul, as is marked with a yellow box in Figure 5-1. This area is thereby known to be dynamic and fast changing, which makes it a good test region for change detection. The images (a1, b1) were captured on the 24th of August 2011 and (a2, b2) at the 7th of July 2012. They have a size of 350 x 400m (a) and 500 x 500m (b) respectively.

Figure 5-1: study area

Figure 5-2: study sites: (a1, b1) Sites at the first point of time. (a2, b2) Sites at the second point of time

All images are from the WorldView-2 satellite, which has one panchromatic band and eight spectral bands. The multispectral bands are covering a spectral range from 400nm to 1050nm and are having a spatial resolution of 1.84m (Padwick et al. 2010). And it has the ability to receive along track stereo data.

The panchromatic band (see Figure 5-2), covers thereby a spectral range from 450nm – 800nm. The DSM was calculated from this panchromatic stereo data and has originally 1 meter resolution (Tian and Reinartz, 2013). They have been resampled to 0.5 meter resolution in order to have the same resolution as panchromatic images. After pansharpening the resolution of the multispectral image could be improved to 0.5m, the same as the resolution of the panchromatic image.

However, as can be seen in Figure 5-4 , the quality of the DSM is sufficient to detect large size buildings; smaller ones are though barely visible. The boundaries of the objects in the image are usually very blurry and due to some drawbacks in the DSM generation, the height values of some objects are not correctly represented. An example for this is the red marked building in Figure 5-2 (a1), which is marked with a black circle in Figure 5-4 (a1, a2). It could clearly be seen that the building does not have the right height in Figure 5-4 a1, while it is correctly represented in Figure 5-4 a2.

5.2 Results

In this section, the results of the four main methods DSM – refinement, segmentation, building footprint extraction and change detection are presented as well as and the best found parameter settings.

5.2.1 DSM refinement

The panchromatic image is adopted for the DSM refinement. Therefore, a preprocessing of the panchromatic image is performed. Firstly, the contrast on the panchromatic image was stretched. This was done by setting the outliers, which are all values more than 1.5 standard deviations apart from the mean value, to defined maximum or minimum values. Then the median filtering was performed with a 5 x 5 window size.

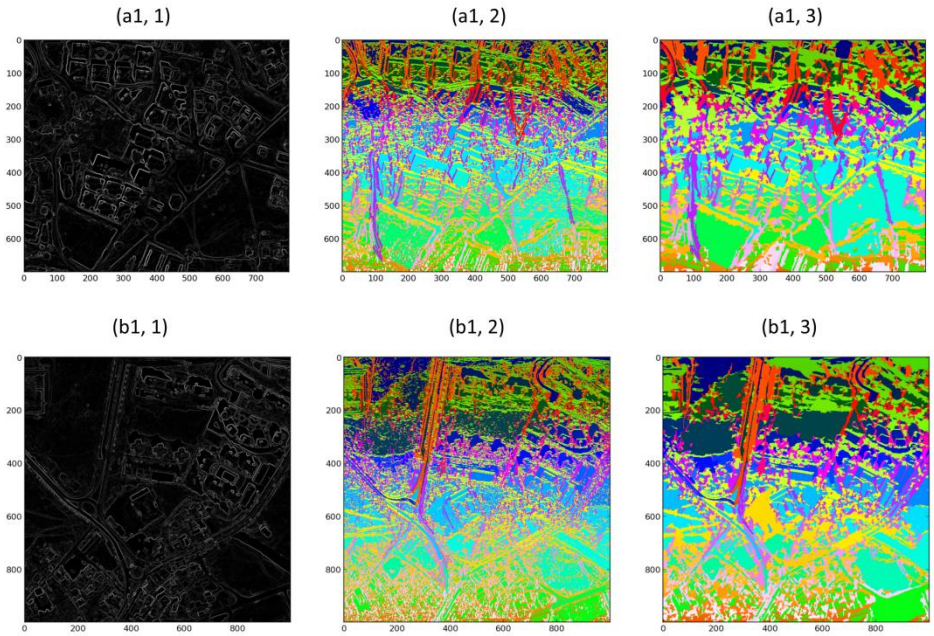

Figure 5-3: Segmentation steps for (a1) and (b1). (1) Sum of the Sobel gradients in horizontal and vertical direction. (2) Initial segmentation. (3) Improved Segmentation due to merging.

For the determination of the boundaries, the Sobel gradients in vertical and horizontal direction were calculated and added up, as is shown in Figure 5-3 (a1, 1) and (b1, 1). Next, a threshold of 0.5 times the average gradient over the whole image was used to distinguish between boundaries and background. The Segmentation result, is shown in Figure 5-3 ((a1, 2), (b1, 2)), and was refined afterwards with merging. Thereby all segments with a size smaller than 10 pixels were assigned a “nodata” – value. The areas of these, which were encapsulated by a larger area by an amount of 75%, are merged to the encapsulating area. The remaining “nodata” – areas got the value of the most common value in their surroundings, starting with a 3x3 window that is enlarged if there are just “nodata” - values within.

After that, triple merging was performed for all segments. For that, a height threshold was set to 2m and an intensity threshold of 15% of the maximum intensity in the image. The middle part of the three segments (M) thereby was defined as a boundary segment with two main neighbours that need to be area segments. In order to be a main neighbour, they need to cover at least 35% of the whole boundary of M each.

The result of this merging procedure is shown in Figure 5-3 ((a1, 3), (b1, 3)). After merging, the number of segments could be heavily decreased to about 10% of the original amount, as is presented in Table 1.

Table 1: Number of segments before and after merging

	Test site a1	Test site a2	Test site b1	Test site b2
Number of segments before merging	21241	20795	36201	37043
Number of segments after merging	2112	2304	3641	3670

The improved segmentation was used then, to calculate the average height over the segments. With the error check method explained in section 4.2.3 it was searched for areas, where the change between the average height and the original DSM was reasonable. This was assumed when the average difference over a segment was smaller than 5% of the average height of the segment. In such cases the height values of pixels in the DSM were updated.

Both, the original and the refined DSM are shown in Figure 5-4. The improvements can be found in detail, especially on the building boundaries. It has to be noted, buildings that are standing close together are getting a better separation. Moreover the boundaries of many objects are sharper than before. Some good examples for these little changes are marked in Figure 5-4 with red circles.

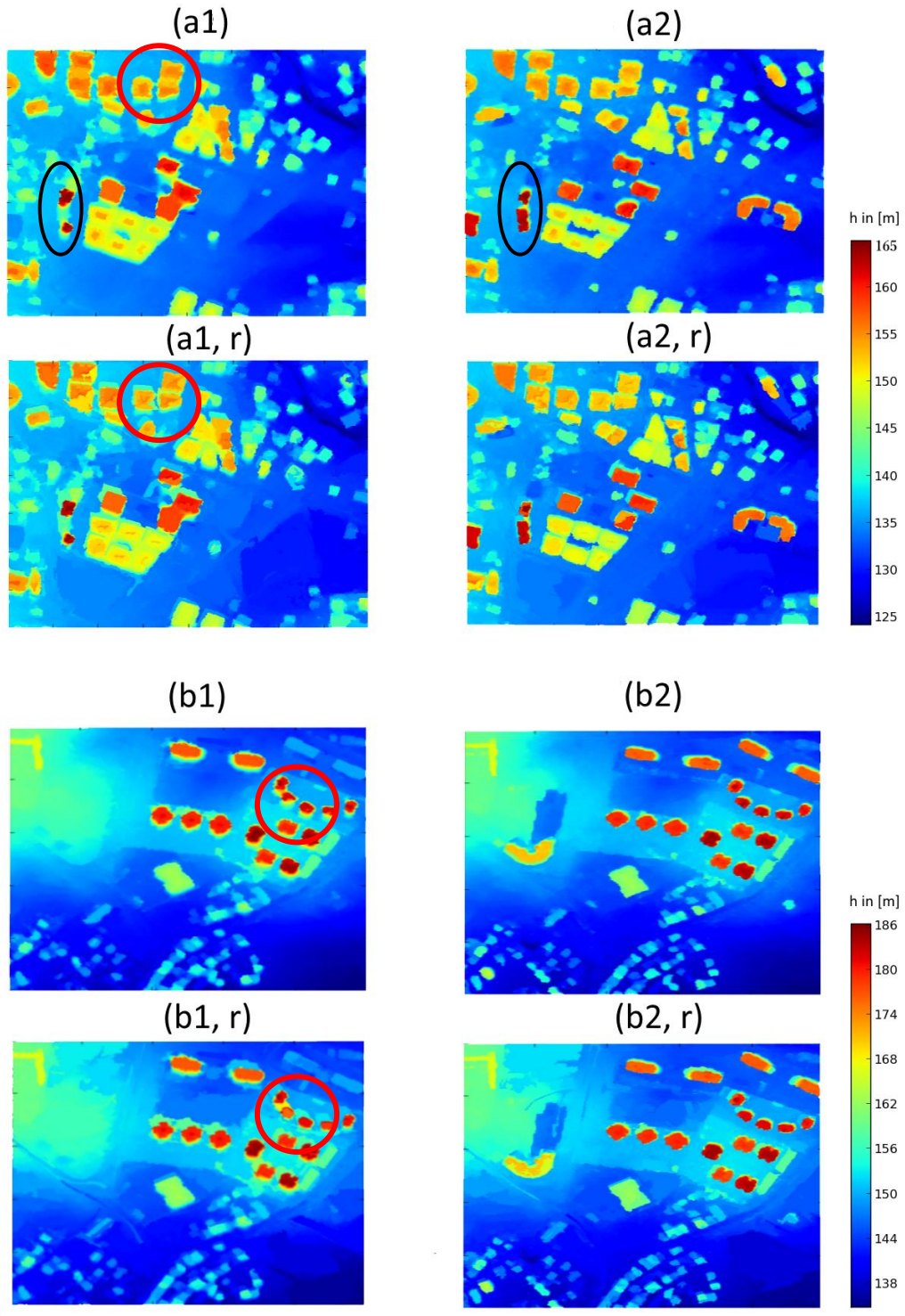

Figure 5-4: Comparison between the original and the refined DSM's. (a1, a2, b1, b2) Original DSM's. (a1r, a2r, b1r, b2r) refined DSM's.

5.2.2 Segmentation

Based on the method illustrated in section 4.3.1, firstly the PCA is performed on the pansharpened multispectral images of all test data. The original multispectral images (a1, b1) and the PCA- results (a1pca, b1pca) can be seen in Figure 5-5. It is clearly visible

(see red marked places), that objects in the image are now easier detectable, and the ground and building roofs can be better separated

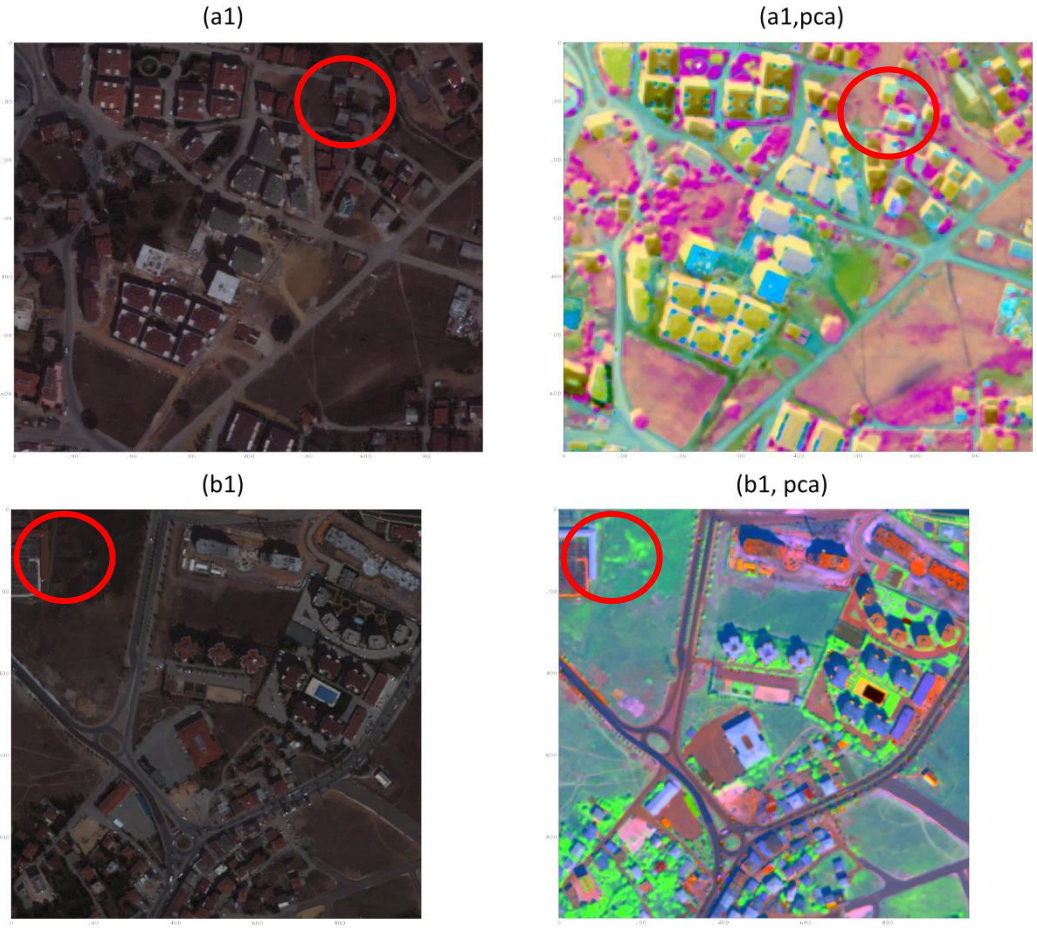

Figure 5-5: Comparison between the multispectral images (a1, b1) (channels 5, 3, 2) and the pca – results (a1pca, b1pca).

Subsequently, the input images need to be converted to feature space. Thereby the range of the images is representing their weight in the mean shift segmentation. The scale ranges of the images are therefore stretched, as it is shown in Table 2. The mean shift segmentation was then performed with a uniform kernel with a size of 30.

Table 2: Scale range of the input features

	PCA – channels	Panchromatic image	Refined DSM	X/Y Position
New scale range	0-200	0-500	0-600	Kept original

The segmentation result was subsequently improved by merging small segments. Thereby all segments smaller than 20 pixels were merged to their neighbour with the smallest height difference. For it is possible that there are merged segments with a small size left, the procedure was repeated three times. The last step is the encapsulation-based check. The threshold for the encapsulation was thereby set to 80%. By doing so, the number of segments decreased as to roughly about one third, as is shown in Table 3 and Figure 5-6.

Table 3: segment reduction due to merging

	a1	a2	b1	b2
Old number of segments before merging	16020	16350	26626	23522
New number of segments after merging	5668	5515	8248	7082

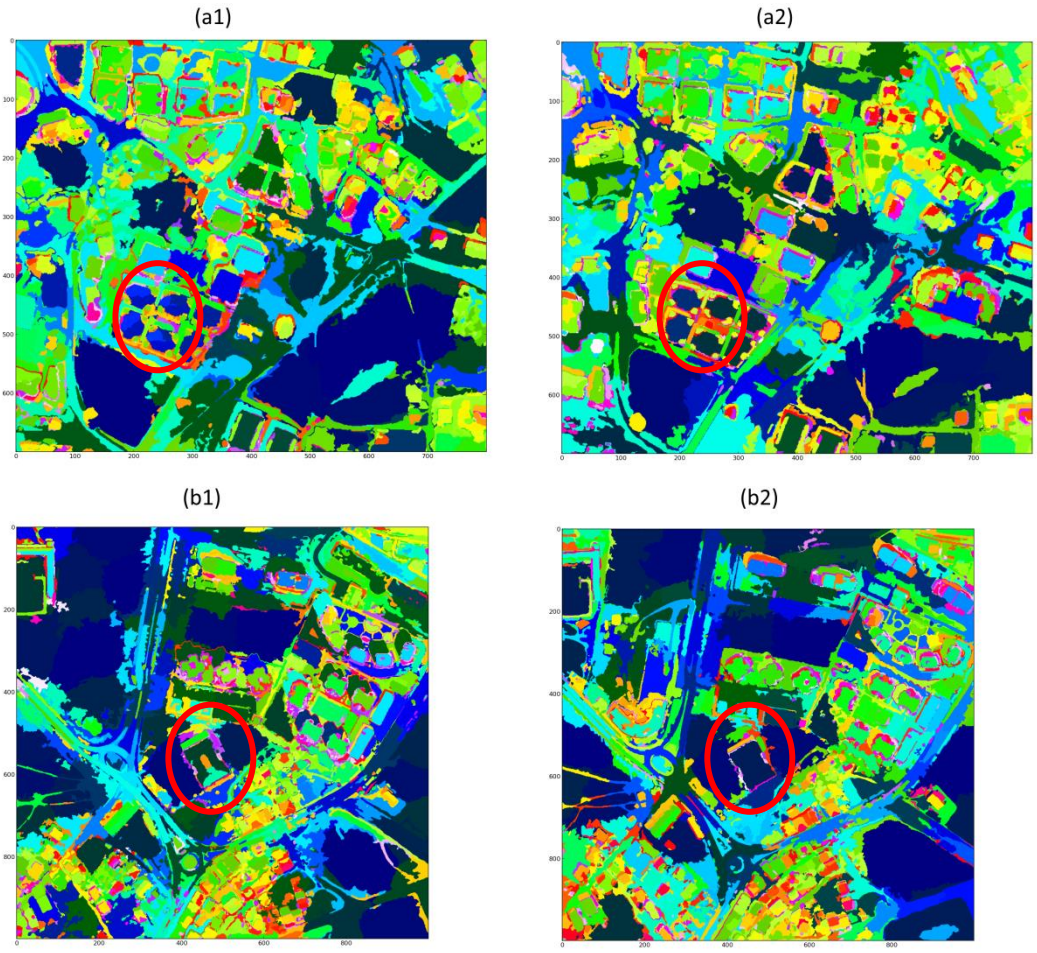

Figure 5-6: Refined segmentation results after merging for the areas a1, a2, b1, and b2.

As can be seen, the segmentation at this point is representing the buildings well and has less small fragments than before. However, there is still oversegmentation, so buildings are often consisting out of many parts, as is marked in Figure 5-6 (a1, b1) with red circles.

5.2.3 Building footprint detection

At the Beginning of the building footprint detection procedure, building candidate areas need to be defined. Therefore all pixels in the nDSM with a height above 4 m were used, Figure 5-7 (a1, b1).

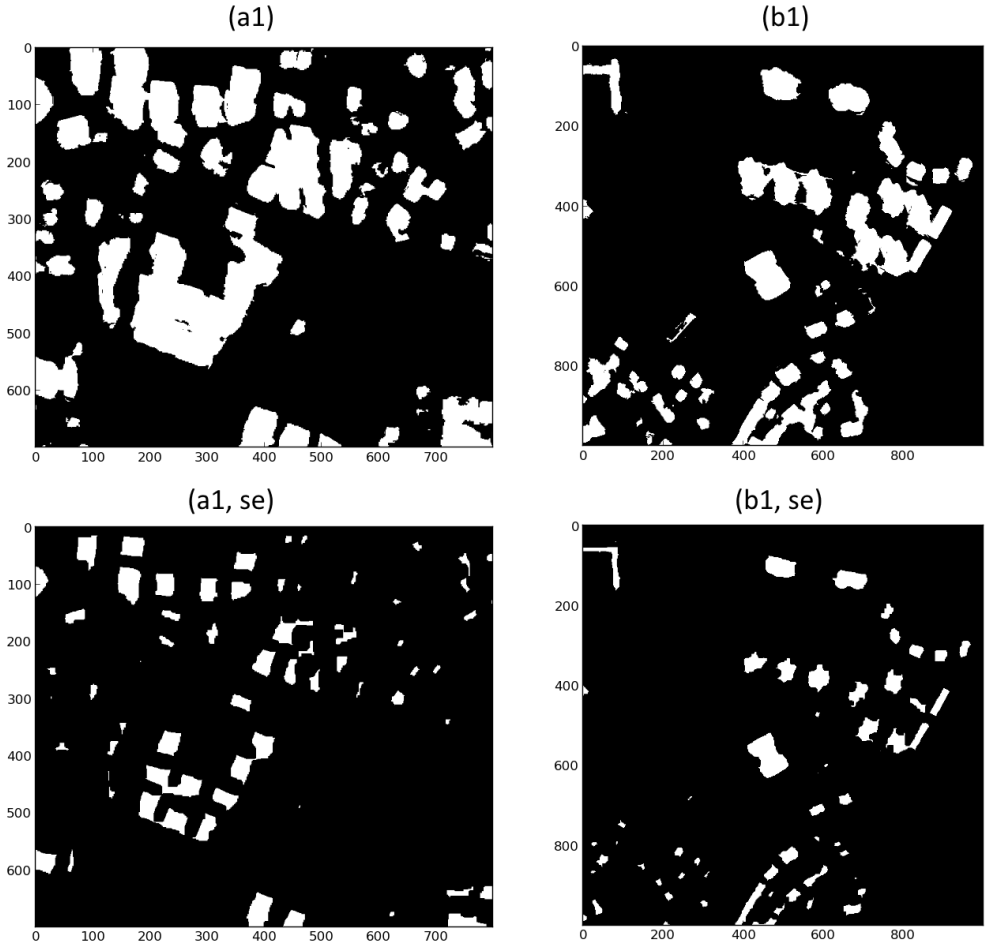

Figure 5-7: (a1, b1) Initial building candidate areas. (a1se, b1se) Building candidate areas after shadow exclusion and erosion.

In order to keep low height buildings, a small threshold value is given. As a consequence some buildings are incorrectly connected together, as shown with red colour circles in Figure 5-7. The results of the following shadow exclusion and erosion ,is shown in Figure 5-7(a1se, b1se)). In this experiment, each segment with a shadow pixel overlay percentage above 35% is excluded from the candidate building regions. Thereby a low threshold is causing a higher risk of losing possible building areas, while a small

threshold might not be sufficient to for the departure. Further separation of the candidate regions is subsequently achieved through erosion by 15 pixels.

The final building candidate regions (shown in Figure 5-7 (a1se, b1se)) are then dilated by the same amount, iterated and the associated segments are determined. Thus, the segments need to be covered by a candidate area by at least 70%.

All of the segments within a candidate region are then tried to merge. Thereby, at first all areas with a size below 70 pixels are merged to their neighbours, if the height difference is smaller than two meters. After that, merging is performed for all segments by using the multi feature merging procedure. The thresholds and weights for the features are thereby shown in Table 4. Two segments were merged, when their sum of the weights was above four.

Table 4: Thresholds and Weights for the multi feature merging

	PCA –Channels 1- 3	Intensity	Height
Threshold	0.2* σ (pca_img)	1* σ(pan_img)	2m
Weight	1	1	2

As next, it is searched for the segments which are part of a building in the candidate region. This is done only in dependency of the height. The local threshold is set to the average height in the region minus 1.5 *the standard deviation in the whole scene. All segments with an average height higher than the threshold are forming the building and are transferred into the building footprint. The building footprints are shown in Figure 5-8 (a1, a2, b1, b2).

5.2.4 Change detection

For change detection two methods were examined. In one of them, the two building footprints of one scene are compared with each other. Thereby it is needed to calculate the overlay of the buildings, as is shown in Figure 5-8 (acov, bcov). The value of a segment is thereby representing the ratio of uncovered pixels (therefore referred to as acov, bcov in the image heading) in the building areas. All Segments with a ratio above 0.7 are considered to be new buildings. The resulting change features using this method can be seen in Figure 5-8 (achange, bchange)

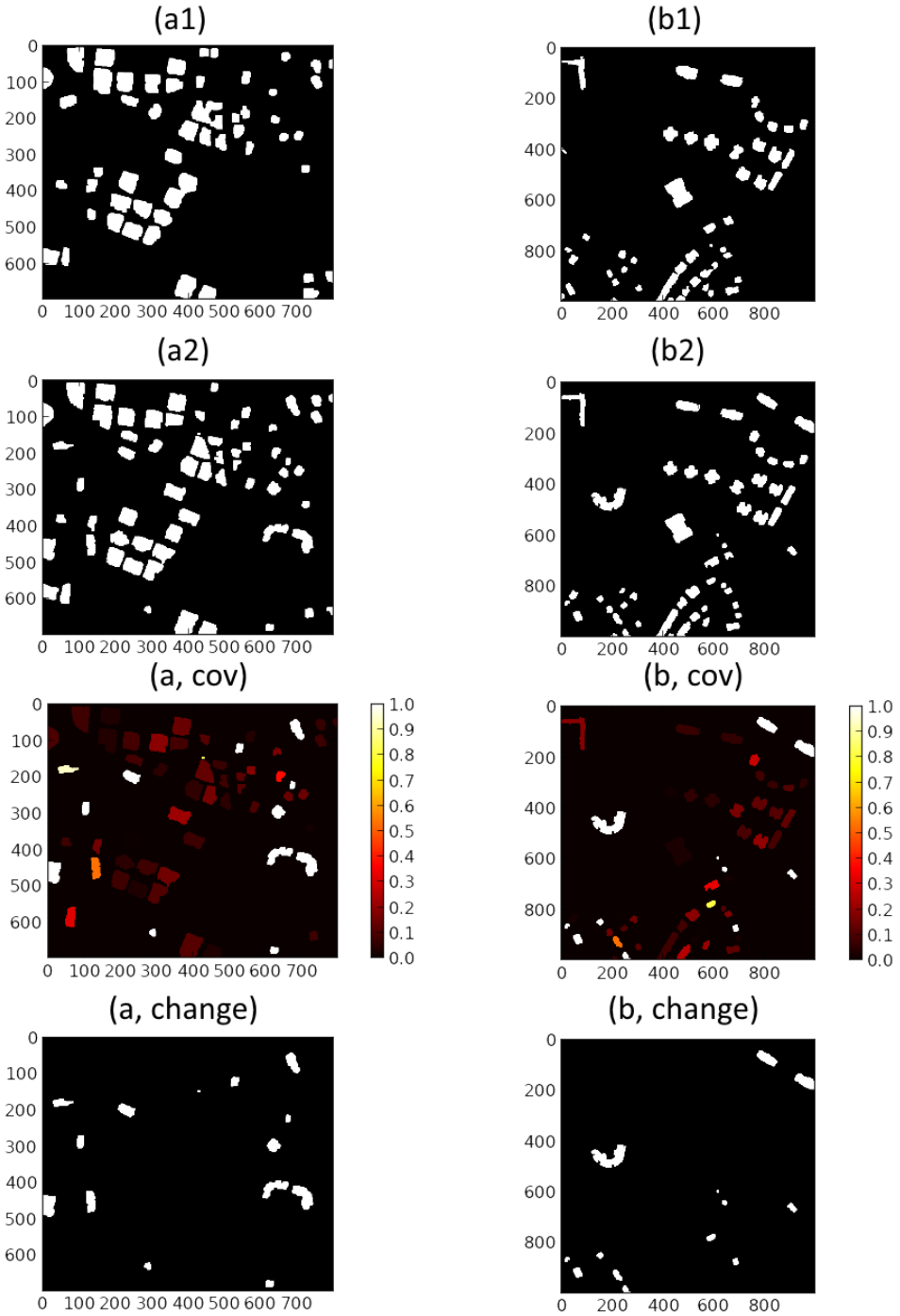

Figure 5-8: Change detection by comparing the building footprints. (a1, b1) Building footprints from the first point of time. (a2, b2) Building footprints from the second point of time. (acov, bcov) Ratio of uncovered pixels in building area. (achange, bchange) Change detection result.

The other approach is to calculate change indicator images first, which are then used together with the second building footprint mask, in order to perform change detection.

The change indicator is thereby representing the change probability for each pixel, as is shown in Figure 5-9 (aci, bci). All pixels with a change probability above 0.3 are then

transferred into a binary image. Again the overlay of segments could be calculated and all buildings with a ratio of uncovered pixels above 0.7 are considered to be new buildings. The resulting change features can be seen in Figure 5-9 (achange, bchange).

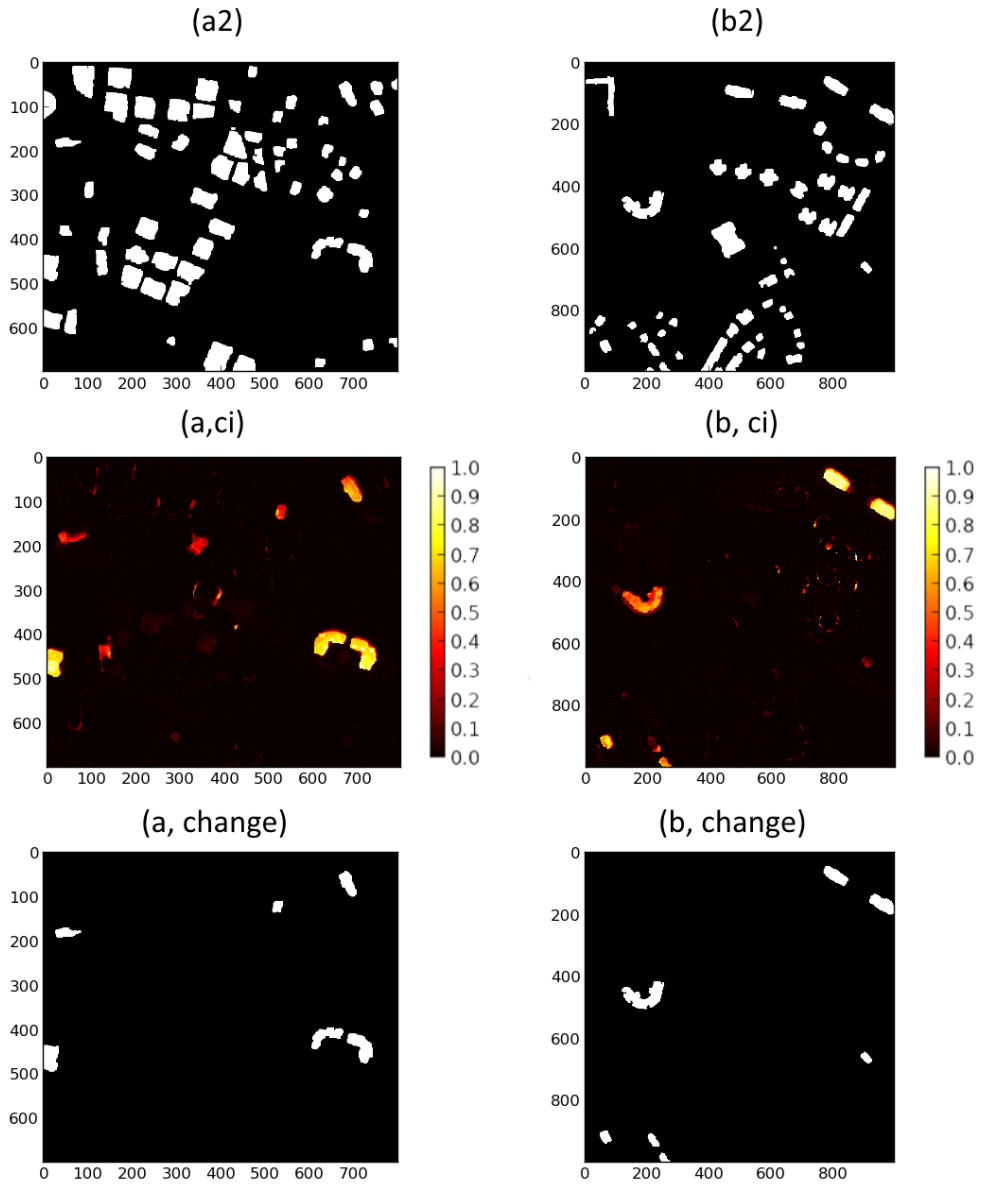

Figure 5-9: Change detection based on building footprints combined with change indicators. (a2, b2) Building footprints from the second point of time. (aci, bci) Change indicator images. (achange, bchange) Change results.

5.3 Evaluation

In this part, the building footprints, as well as the change detection results are evaluated.

5.3.1 Evaluation of the building footprints

For the evaluation, the calculated building footprints (a2, b2) are compared to the reference building footprints (see Figure 5-10 (a2ref, b2ref)).

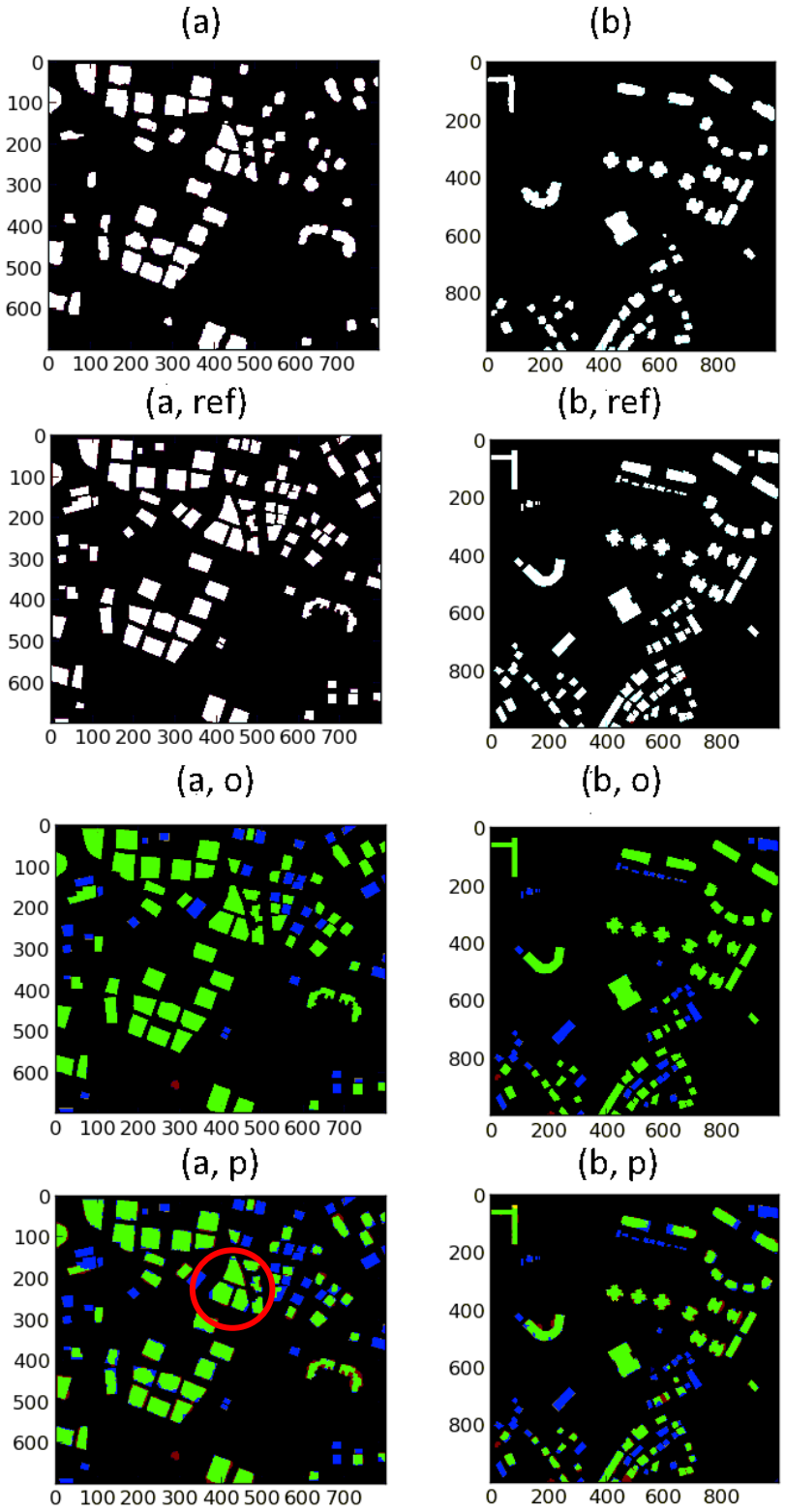

Figure 5-10: Evaluation of the building footprints. (a2, b2) Building footprints. (a2ref, b2ref) Reference building footprints. (a2o, b2o) Object based evaluation. (a2p, b2p) Pixel based evaluation.

Green: true positive, Blue: false negative, Red: false positive, Black: true negative

Figure 5-10 (a, o) and (b,o) are the object-based evaluation results. An object in the result is detected as true positive if it has an overlay more than 50% of the ground truth.

The evaluation results of the object-based evaluation are recorded in Table 5, where the number of true- and not detected objects, as well as the number of false alarms is presented in dependency on the object size. Further Figure 5-10 (a2o, b2o) is showing the results graphically. Thereby objects can be classified as true positive (green), false negative (blue), false alarm (red), or background (black). True positive is thereby meaning that building regions/ pixels are correct detected as buildings. False negative is the case, when building regions/pixels are detected as no-building. The opposite of that is false positive, where no-building regions/pixels are falsely detected as buildings. The true negative class enfolds all pixels that are background, both, in the calculated and the reference image.

As can be seen in Figure 5-10, the majority of the large size buildings are green and therefore right detected, while especially small buildings are often blue. This trend is also represented in Table 5. While buildings with a size smaller than 125 [m^2] are just in a few cases detected, buildings larger than 375[m^2] are almost all detected. False positive objects were thereby not often recognized.

Table 5: Object based evaluation of the building footprints.

	building size < 125 [m^2]		125[m^2] < building size < 250[m^2]		250[m^2] < building size < 375[m^2]		building size >= 375[m^2]		Over all
	True	False negative	True	False negative	True	False negative	True	False negative	False positive
a1	1	20	21	10	11	4	19	0	0
a2	7	24	20	9	11	2	9	0	1
b1	4	17	19	17	9	2	15	4	2
b2	9	30	18	9	8	1	11	2	2

Besides the object-based approach, pixel based evaluation is shown in Figure 5-10 (a2p, b2p) and Table 6, where it can be seen how accurate the boundaries of the detected objects are determined. While some boundaries are matched well, in other buildings (see red marked object) there are missing some parts. The referring statistics are presented in Table 6, where the number of pixels for each of the classes can be seen, for all four test regions.

Table 6: Pixel based building footprint evaluation.

	True positive	False negative	False positive	True negative	KIA
a1	65434	30476	8218	455872	0.73
a2	68818	24781	9385	457016	0.77
b1	59355	46066	11198	883381	0.64
b2	70056	32301	11282	885871	0.74

As well as the pixel based statistics, as is shown in Table 6, the kappa index of agreement (KIA) is calculated, which is a measurement for the quality of the results (Cohen 1960). It is calculated as presented in equation (22). Thereby, $\Pr(a)$ is the relative agreement, which is calculated in the case above, by adding the number of true detected pixels (TP) with the number of true not detected pixels (TN). This sum is then divided by the total number of pixels in the image (N). $\Pr(e)$ is the probability of random agreement. It is calculated by adding the probability of a pixel randomly being detected as a building in both images, and the probability of a pixel being randomly detected as no building in both images (Cohen 1960). The calculation method is shown in equation (21), thereby (FN) stands for false negative and (FP) for false positive.

(21)

$$\Pr(a) = \frac{TP + TN}{N}$$

$$\Pr(e) = \frac{(TP + FP) * (TP + FN) + (FN + TN) * (FP + TN)}{N * N}$$

(22)

$$KIA = \frac{\Pr(a) - \Pr(e)}{1 - \Pr(e)}$$

As is presented in Table 6, the KIA is in a range between 0.64 and 0.77. Thereby in both scenes, the images from the after-event data get a higher KIA.

5.3.2 Evaluation of the Change detection results

The region/ pixel based building mask evaluation approaches are further used to evaluate the change detection results.

Figure 5-11 thereby shows the evaluation for the post-classification building change detection method. It can be seen (ao, bo) that again the most of the larger new buildings are detected, while especially smaller ones are missing. However, there is also a big number of false alarms (red objects) that is leading to smaller KIA values (0.51 for test site a and 0.65 for test site b, see Table 7) compared to the other approaches.

Figure 5-12 is showing the evaluation of the object building change detection method, which is using the building footprints as well as change indicators. At that, the true detected objects are the same as those from the building footprint comparison method (see ao, bo in Figure 5-11 and Figure 5-12). This is also the case for the false negative class. Though, a big difference can be seen in terms of false alarms. While there are ten (see Table 7,Table 8) by using the first method, there are just two in the object building change detection approach. This is also reflected by the KIA (see Table 7), which is in the second approach 0.66 for scene a, and 0.71 for scene b.

The evaluation of the just pixel based method is shown in Figure 5-13. It can clearly be seen, that this approach is producing a lot of false alarms. Because the original building footprint masks has already ten false positive objects in total, this one has twenty (see Table 8). However, only this method is able to detect the red marked building in Figure 5-11, Figure 5-12 and Figure 5-13. The other buildings that were not detected by the two other approaches are not detected by this method as well. As can be seen in Table 7, the KIA is for both scenes 0.63. For that, it is a bit better than the method based on building footprint comparison, but not as good as the object based method that uses the change indicator and the building footprints.

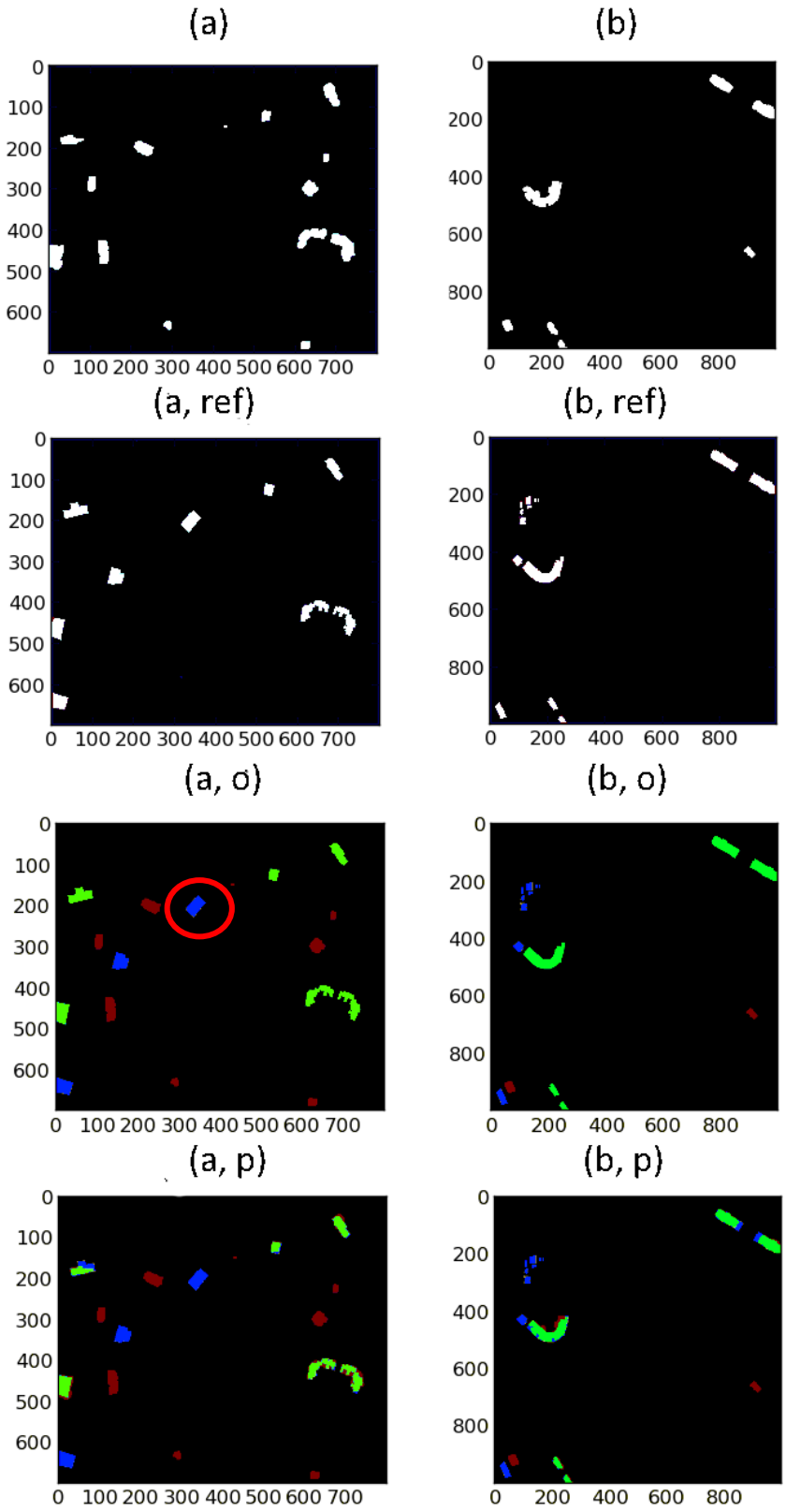

Figure 5-11: Evaluation of the change detection results based on building footprint comparison. (a,b) Change detection results. (aref, bref) Reference images. (ao, bo) Object based evaluation. (ap, bp) Pixel based evaluation.

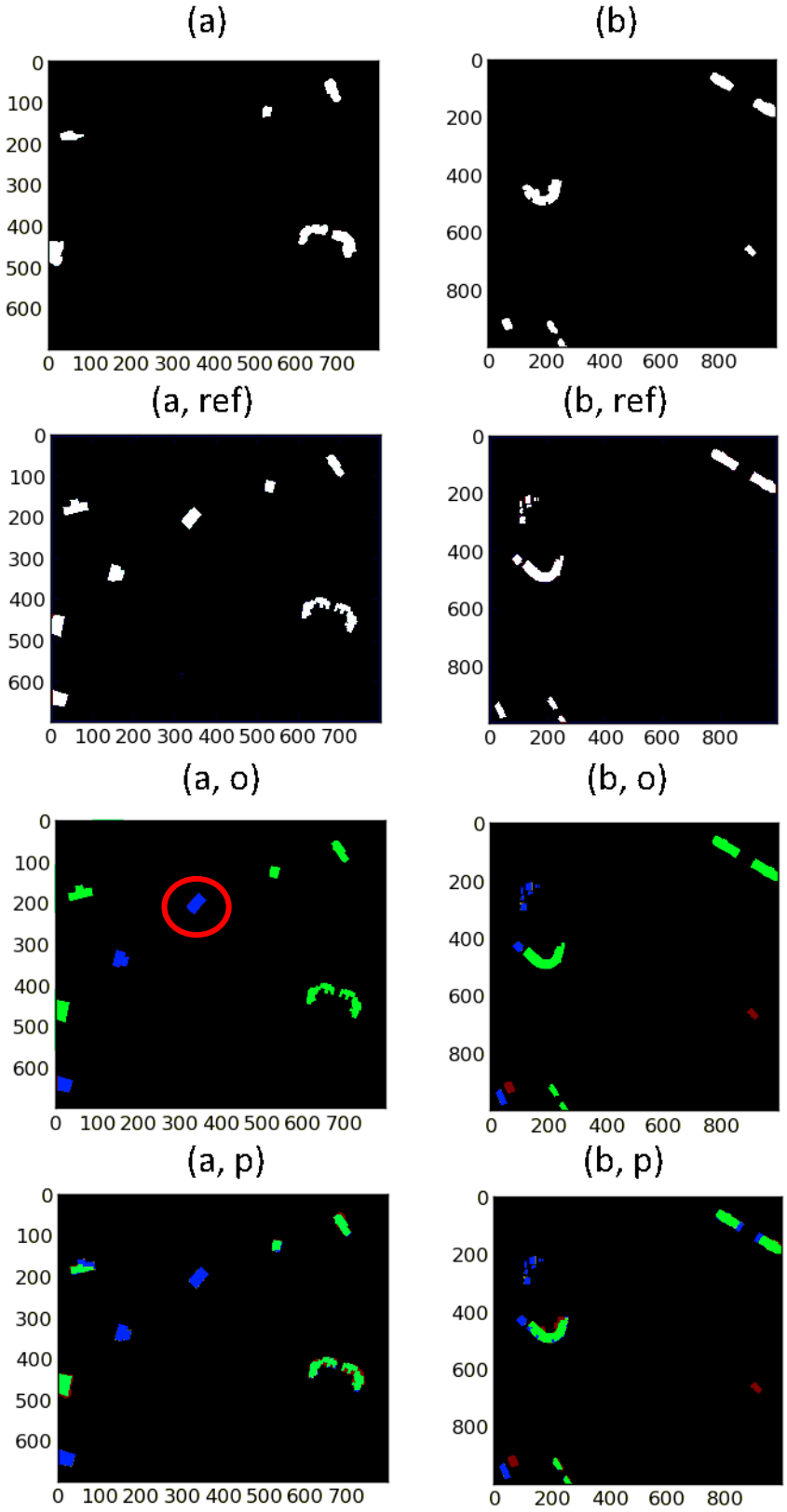

Figure 5-12: Evaluation of the change detection results based on building footprints combined with change indicators. (a,b) Change detection results. (aref, bref) Reference images. (ao, bo) Object based evaluation. (ap, bp) Pixel based evaluation.

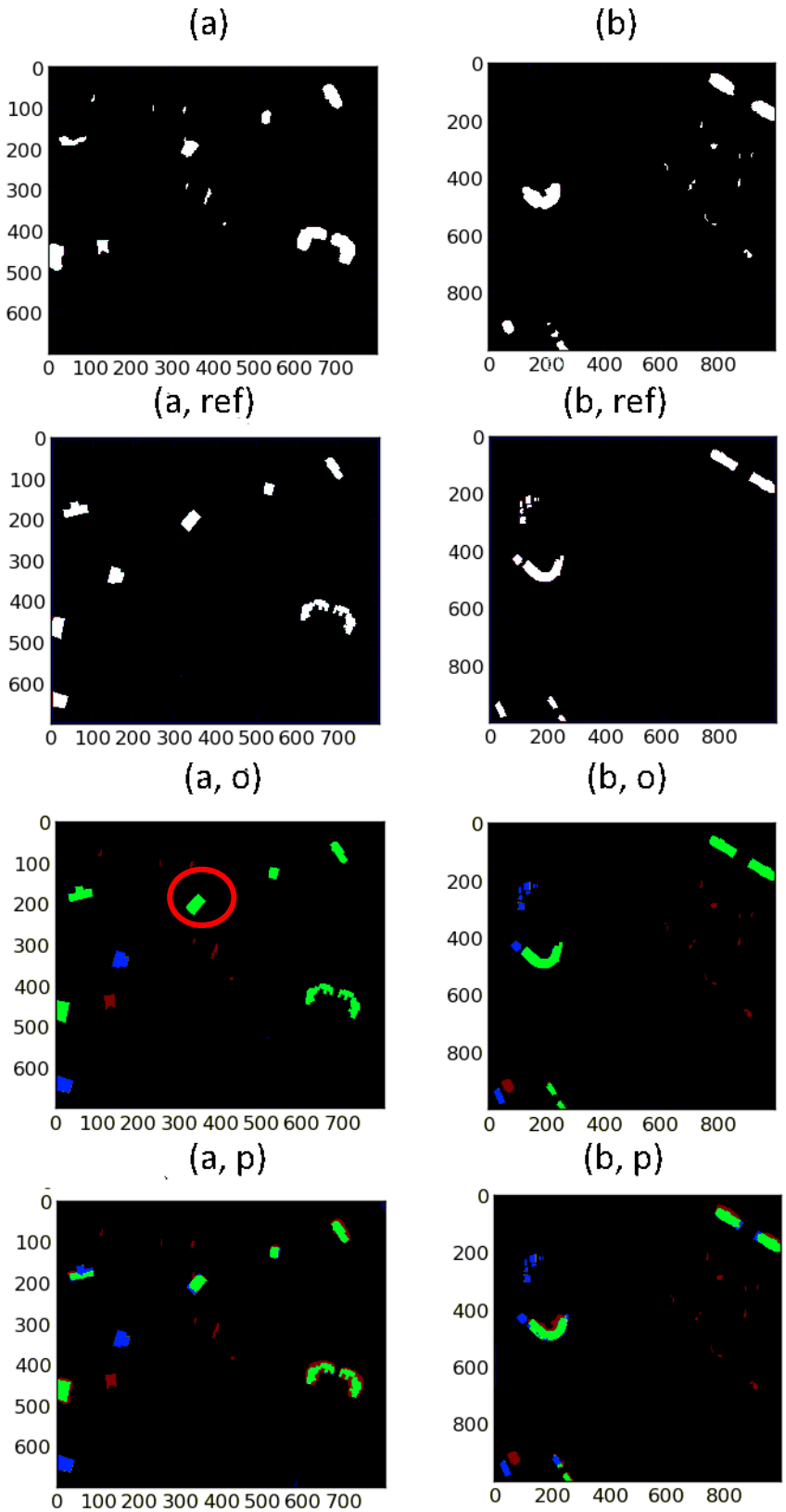

Figure 5-13: Evaluation of the change detection results based just on change indicators. (a,b) Change detection results. (aref, bref) Reference images. (ao, bo) Object based evaluation. (ap, bp) Pixel based evaluation.

Table 7: Pixel based evaluation of change detection results for the images a, b based on: (fp) building footprint comparison, (fc) building footprints and change indicator, (c) just change indicator

	True positive	False negative	False positive	True negative	KIA
fp - a	6485	4653	7211	541651	0.51
fp -b	12136	6032	6415	975417	0.65
fc –a	6485	4653	1993	546869	0.66
fc - b	12136	6032	3840	977992	0.71
c - a	7736	3402	5483	543379	0.63
c - b	12083	6085	7624	974208	0.63

Table 8: Object based evaluation of change detection results for the images a, b based on: (fp) building footprint comparison, (fc) building footprints and change indicator, (c) just change indicator

	True positive	False negative	False positive
fp - a	6	3	8
fp -b	5	9	2
fc –a	6	3	0
fc - b	5	9	2
c - a	7	2	7
c - b	5	9	13

6 Discussion

Subsequently the results of the four main steps should be analysed and discussed.

6.1 DSM refinement

In 5.2.1 it is demonstrated that the DSM refinement method is leading especially to a better distinction of buildings that are standing close together. Further, single objects are getting a more homogeneous height value.

Although the improvements are very limited and just barely visible, they are bringing benefits for the following segmentation and are so improving the final results. Segmentation without the refinement would be also possible, but the results would be more vulnerable to errors.

In order to accomplish the DSM refinement, many single methods are used. Especially the merging procedures thereby need a lot of computation time (round about 2min for a), while the sobel gradient based segmentation itself is taking just a few (~5) seconds. For that, it might be considered to shorten the merging, and with that some of the improvement, for a better performance.

The segmentation based on the sobel gradients thereby proved to be reliable, although only the panchromatic image is used as input. In cases where the boundaries of buildings were not sharp in the panchromatic image, it can however happen that objects are not delineated from each other, as is shown in Figure 4-5. Other segmentation procedures, that are using multiple features like the mean shift segmentation, are more robust for such problems. However, the mean shift segmentation needs much longer computation time (~2min for a) for the processing. Since it is possible to eliminate those errors with the change check at the end, the faster sobel gradient segmentation is used.

Another advantage of this segmentation method is that pixels belonging to boundaries/ high gradient areas are well recorded. This information can be further used in the merging, as in the height assignment in the end, where those boundaries areas are not used for height changes and the accuracy of the refined DSM is higher, as only clear separated homogeneous areas are assigned the average value over the segment.

As can be seen in Figure 4-13, the DSM refinement leads to a better matching of the building shapes. However, the new height is just an average value and might therefore not represent the actual building height. Especially in cases like shown in Figure 4-13, where large parts of the buildings are not recognized in the original DSM, as a future work, a new method is required, which can estimate the real height of these buildings. Therefore, the frequency distribution of those areas can be examined and the new height might be dependent on the most common values.

6.2 Segmentation

The segmentation is needed for the following building footprint extraction. In order to match of the buildings well in the building footprint, it is important that the segmentation is

as precise as possible. Oversegmentation is thereby causing a higher risk of missing building parts, while undersegmentation is leading to inaccurate building boundaries, where multiple buildings might be detected as one.

Oversegmentation could be handled with merging, undersegmentation is more difficult to refine. Therefore, the first segmentation results are oversegmented and later refined. For this, it is important to have a first segmentation that is matching the object boundaries well. For those boundaries are sometimes not well recognizable in some of the images, multiple input features are used. Despite this is leading to longer processing times, this needs to be done in order to get good results.

To reduce oversegmentation, merging is performed for the whole image, although just the building areas are of interest. This again is leading to long processing times. For there is in the next method step - the building footprint extraction- building based merging performed, it needs to be considered to leave out the merging at this step. Though, this would lead to small segments sometimes not being recognized as parts of the candidate regions, which is induced by the inaccurate boundaries of the building candidate regions. Small segments at these boundaries might not have a sufficient overlay percentage with it and would therefore be not recognized as a candidate building segment. This would cause a minor increase of the inaccuracy of the extracted building footprints, which would be traded for a better speed of the method.

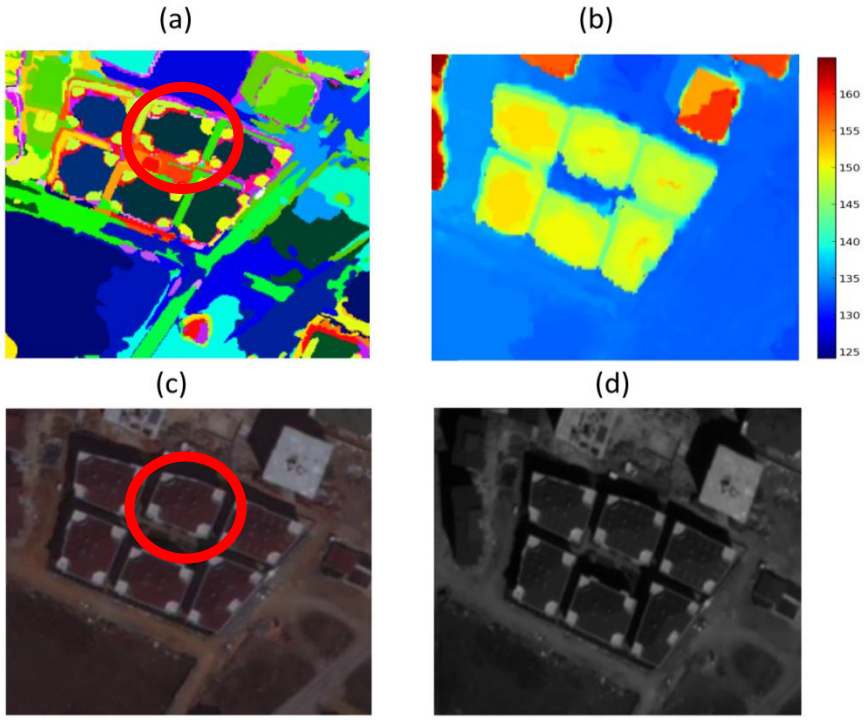

Figure 6-1: Unmergeable segments. (a) Segmentation. (b) Refined DSM. (c) Multispectral image. (d) panchromatic image.

The refined segmentations, as shown in Figure 5-6, are representing the buildings well. Some buildings are however still consisting out of many parts. This is due to large differences within the segments of a building, which can be seen in marked with a red circle in Figure 6-1. At this marked building, the corners of the building are very different from the middle part. While some times the multispectral channels can help to find similarities between the segments, at that building, all input channels are indicating an object change. For that, it is not possible to merge these together, as using a lower threshold would cause undersegmentation errors at other places. Concluding, oversegmentation cannot be completely eliminated from the segmentation.

6.3 Building footprint extraction

For building footprint extraction, the nDSM needs to be calculated first. This is then used to get an initial building candidate mask, as is shown in Figure 5-7. In that step, a global threshold is used, which leads to a barely accurate candidate mask. It might be considered to refine this mask first, before shadow exclusion and erosion is performed. This can be achieved by using a local height based threshold already at this point, for the height differences within these areas are large enough for a separation.

As next, shadows were excluded from the mask. While the most shadow regions are useful for the separation of objects, it could also happen that buildings are covered with shadows or being mis-detected as shadows. Thus they are removed from the building candidate mask. This is for example the case with the red marked building in Figure 6-2, which is concluding is not represented in the final building candidate mask.

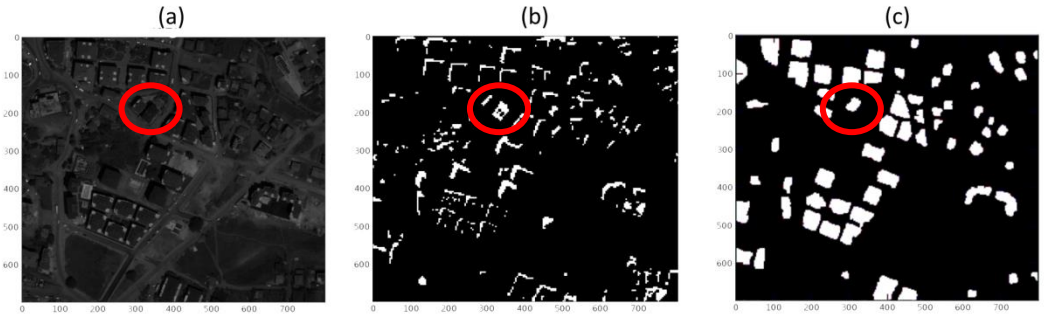

Figure 6-2: Building loss due to shadow exclusion. (a) Panchromatic image. (b) Shadow mask. (c) Building footprint.

After the shadow exclusion, erosion is performed to further distinguish single objects. Indeed, as is shown in Figure 5-7 it is effective, but using the erosion is also causing the complete elimination of small objects in the candidate mask. For that, a good balance

needs to be found between a sufficient separation of buildings, and keeping also small objects.

If the objects would be better separated before, like possible with the before made proposal, a smaller erosion would be useable. Further, it might be tried to eliminate the vegetation, like proposed by Gamba and Houshmand (2000). Especially in residential areas with gardens, this would help to separate buildings from their surroundings and to eliminate vegetation from the building footprint. This would cause a smaller number of false alarms in building footprints.

After the building candidate mask is determined, the buildings in these candidate areas are searched. This is done with a local threshold just based on height. Thereby it might be considered to use multispectral information as well. Like proposed by (Lee et al. 2003) it is also possible to classify building roofs. This information can be used to further improve building detection results.

In this approach however, all segments with a height larger than the local height based threshold are assumed to be buildings. Thus, it is possible to detect and separate multiple buildings within one candidate area. However, sometimes there are parts of the buildings missing, like is shown in the red marked area in Figure 5-10. This is due to height differences within buildings and due to the already discussed cases of oversegmentation in buildings, which cannot be fully eliminated. Other approaches, like (Lee et al. 2003) and (Tian and Reinartz, 2013) propose to use Hough lines to extract accurate building shapes. Hough lines are thereby representing the main lines in an image. In case of having just one, or a few quadrangle buildings within one candidate area, this might truly increase the accuracy of the building footprints. Having many buildings in one candidate area, might however decrease the accuracy, for to many main/Hough lines are causing again oversegmentation.

Additional to these method based errors, some buildings could not be detected for errors in the data. Especially the inaccuracy of the DSM is thereby causing some missing buildings, for objects need to have a certain height in order to be determined as a building.

Overall, it can be stated that it is possible with this method to detect the majority of the big buildings, while small buildings are not well recognized. Due to the use of the previous segmentation, the shapes of the building footprints are close to those of the real buildings. Sometimes however, small parts of the buildings are missing, which is decreasing the accuracy.

6.4 Change detection

For Change detection two methods were proposed. Thereby one was based on the comparison of building footprints. The results of this method showed to be more

inaccurate than the other evaluated ones. While the number of true positive and false negative objects were almost the same compared to the other approaches, the large number of false positive objects was leading to a smaller KIA.

This is caused by the summed up inaccuracy of both building footprints. Each building that is missing in the first building footprint thereby will be detected as a changed object, as long as it is not also missing in the second building footprint. Vice versa, buildings which are missing in the second footprint cannot be detected as a change objects.

Further, it needs to be considered that the processing takes much longer time, as both the pre- and after event images have to be processed. However, if there would be already one building footprint available - for instance some GIS cadastre data etc.- this approach might be useful and more efficient, to detect change objects. Also the number of false positive objects would decrease drastically, for having an ideal comparison base. This would lead to a better accuracy.

The second method, which is based on the second date building footprints and the change indicator images, proved to be the more accurate approach in the evaluation. However, there are several false negatives. One reason for this is the before mentioned problem of missing buildings in the building footprints. As long as a building is not in the building footprint of the second date, there is no chance of detecting them as change objects. For that, the in Figure 6-2 red marked building is missing in the change detection result, despite being recognized in the change indicator image.

The other missing buildings can be ascribed to too low change indicator values. Thereby, height change has a big influence on the resulting change indicator (Tian et al., 2015). The panchromatic images are used as well for the change indicator generation; however, the spectral change is not good to use separately. An example for this is shown marked with red in Figure 6-3. Despite the intensity values of the marked building changed from white to black, there is no real building change. For that, the spectral change is used always in combination with the height change. When there is a spectral change with no height change, the pixels are classified as surface change, which is correct in scenes like shown in Figure 6-3. However, if there is a real change that is not leading to a height difference in the DSM's, the change indicator map will not represent the change. Again the lack of accuracy of the DSM is therefore the reason for not detected buildings.

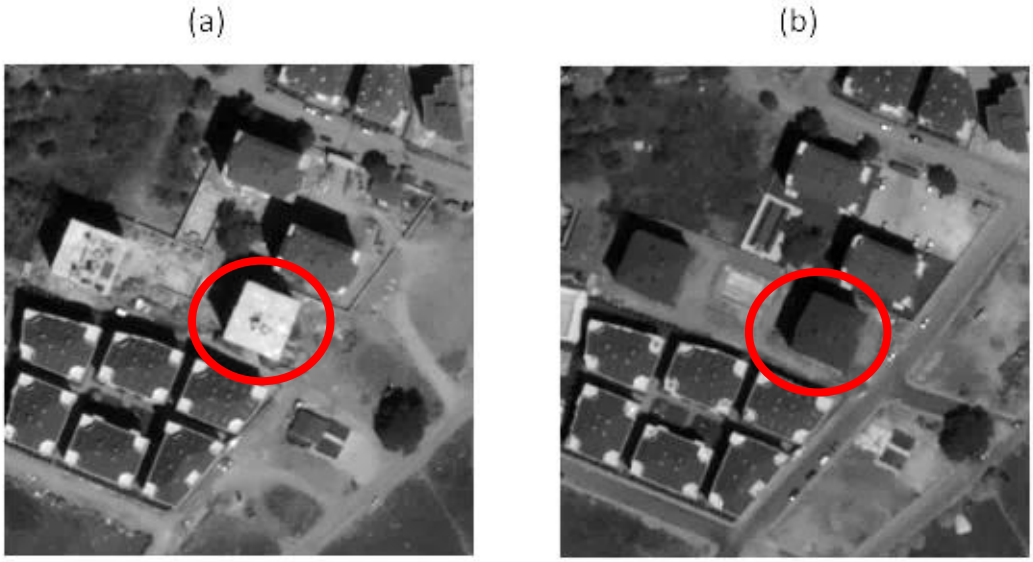

Figure 6-3: Changing reflection of roofs. (a,b) Panchromatic images from one scene.

Using the change indicator on its own is leading to a lot of false positive objects. Further the boundaries of the objects are not as precise as in the combined approach. For that, using both, the building footprint and the change indicator image together seems to be a good method. It might also be considered to use the change indicator together with the segmentation, as proposed by Tian et al. (2013). Thereby the building footprints might be just used for the merging of buildings, while the rest of the image is unchanged.

7 Conclusion

Clearly it was shown that good automatic building footprint extraction and building change detection results can be achieved using WorldView-2 stereo data.

Therefore a method consisting out of four parts was proposed. Thereby, at first the DSM is refined, based on a segmentation attained from the sobel gradients of the panchromatic image. Thus, especially buildings that were standing close together were better separated in the DSM and the boundaries of buildings were sharpened.

Subsequently the refined DSM, the panchromatic image and the multispectral image are used as input features for mean shift segmentation. The segmentation result was thereby representing the building boundaries well and a lot of buildings were recognized as homogeneous segments.

Next, the refined DSM is used to calculate the nDSM, out of which the initial building candidate regions are extracted. These candidate regions are refined, by excluding shadow regions and using morphological opening. Following this, all the segments from

the mean shift segmentation within a building candidate region are determined. Based on the average height of these candidate segments buildings are extracted. The results thereby showed, that the used method works well for the detection of large buildings, small ones are however not this well detectable. Therefore, it is proposed, that future work should focus on the development of better building candidate area refinement methods, which are more sensitive to small building candidate areas.

Three methods are proposed for building change detection. The evaluation proved that the object based approach delivered better results than the pixel based method. It was shown that using a pixel based change indicator together with the second date building footprint delivers better results. Buildings that are not represented in the building footprint though cannot be detected as change objects.

It was shown that the results are very dependent on the DSM, using a more accurate DSM would lead to better results. Together with a further improved method this could lead to better usability of the automatic building footprints and change detection results

As lot of buildings/changes are correctly extracted with the proposed approach, it can be used to get a fast overview over an area. For especially smaller buildings are missing sometimes and the building boundaries are not always very accurate, it should however not be used independently to update maps or databases.

Over all, this work can be seen as a contribution to the building extraction and change detection research. A full application for commercial purposes is yet not reached.

References

Adbi H., Williams L.J. (2010): Principal component analysis. In: Wiley Interdisciplinary Reviews: Computational Statistics, 2 (11), S. 433 – 459.

Blaschke T. (2010): Object based image analysis for remote sensing. In: ISPRS Journal of Photogrammetry and Remote Sensing, vol. 65, no. 1, S. 2-16.

Chen J., Gong P., He C., Pu R., Shi, P. (2003): Land-use / land-cover change detection using improved change–vector analysis. In: Photogrammetric Engineering and Remote Sensing, vol. 69, no. 4, S. 369–379.

Cohen J. (1960): A coefficient of agreement for nominal scales. In: Educational and Psychological Measurement 20 (1), S. 37 – 46.

Comanicu D., Meer P. (2002): Mean shift: a robust approach toward features space analysis. In: IEEE Transactions on Pattern analysis and Machine Intelligence 24(5), S.603-619.

de Miranda P.A.V., Falcao A.X., Udupa J.K. (2010): Synergistic arc-weight estimation for interactive image segmentation using graphs. In: Computer Vision and Image Understanding ,114 (2010), S. 85-99.

Durieux L., Lagabrielle E., Nelson A. (2008): A method for monitoring building construction in urban sprawl areas using object-based analysis of Spot 5 images and existing GIS data. In: ISPRS Journal of Photogrammetry and Remote Sensing, vol. 63, no, 4, S. 399-408.

Gamba P., Houshmand B. (2000): Digital Surface Models and Building Extraction: A Comparison of IFSAR and LIDAR Data. In: IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, Vol 38, No. 4, July 2000, S. 1959- 1968.

Gueguen L., Soille P., Pesaresi M. (2011): Change detection based on information measure. In: IEEE Transactions on Geoscience and Remote Sensing, vol. 49, no. 11, S. 4503-4515.

Hayes D.J., Sader S.A. (2001): Comparison of change-detection techniques for monitoring tropical forest clearing and vegetation regrowth in a time series. In: Photogrammetric Engineering and Remote Sensing, vol. 67, no. 9, S. 1067-1075.

Lee S. D., Shan J., Bethel J S. (2003): Class-Guided Building Extraction from Ikonos Imagery. In: Photogrammetric Engineering & Remote Sensing Vol.69, No. 2, February 2003, S. 143 – 150.

Lu Y. H., Trinder J., Kubik K. (2002): AUTOMATIC BUILDING EXTRACTION FOR 3D TERRAIN RECONSTRUCTION USING INTERPRETATION TECHNIQUES. http://www.ipi.uni-hannover.de/fileadmin/institut/pdf/yihuilu.pdf (01.04.2015)

Lefevre s., Weber J., Sheeren D. (2007): Automatic Building Extraction in VHR Images Using Advanced Morphological Operators. In: Urban Remote Sensing Joint Event, Paris, France, April 11 – 13, 2007, S. 1-5.

Liu H., Zhou Q. (2004): Accuracy analysis of remote sensing change detection by rule-based rationality evaluation with post-classification comparison. In: International Journal of Remote Sensing, vol. 25, no. 5, S. 1037-1050.

Muller S., Zaum D. (2005): Robust Building Detection in Aerial Images. In: International Archieves of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. 36.

Nielsen A. A., Conradsen K., Simpson J. J. (1998): Multivariate alteration detection (MAD) and MAF postprocessing in multispectral, bitemporal image data. In: New approaches to change detection studies. Remote Sensing of Environment, vol. 64, no. 1, S. 1-19.

Ning J., Zhang L., Zhang D., Wu C. (2010): Interactive image segmentation by maximal similarity based region merging. In: Pattern Recognition 43 (2010), S. 445 – 456.

Padwick C., Deskevich M., Pacifici F., Smallwood S. (2010): WORLDVIEW-2 PAN-SHARPENING. From*: ASPRS 2010 Annual Conference San Diego, California, April 26-30.2010.*
www.asprs.org/a/publications/proceedings/sandiego2010/sandiego10/Padwick.pdf
(02.04. 215)

Rottensteiner F., Trinder J., Clode S., Kubik K. (2003): Building Detection Using LIDAR Data and Multi-spectral Images. In: Proc. VIIth Digital Image Computing: Techniques and Applications.Sun C., Talbot H., Ourselin S. Adriaansen [Hrsg.], 10-12 Dec. 2003, Sydney, S. 673 – 681.

San K.D., Turker M. (2006): AUTOMATIC BUILDING DETECTION AND DELINEATION FROM HIGH RESOLUTION SPACE IMAGES USING MODEL-BASED APPROACH. http://www.isprs.org/proceedings/xxxvi/1-w41/makaleler/San_Building_Detection.pdf(27.03.2015)

Scharr, H. (2000): Optimale Operatoren in der Digitalen Bildverarbeitung. Dissertation: Ruprecht-Karls-Universität Heidelberg.

Serra J. (1994): Morphological filtering: An overview. In: Signal Processing. Mathematical Morphology and its Applications to Signal Processing, Vol.38, July 1994, S. 3 - 11.

Shorter N., Kasparis T. (2009): Automatic Vegetation Identification and Building Detection from a Single Nadir Aerial Image. In: Remote Sensing, 2009, 1, S. 731-757.

Tian J. (2013): 3D change detection from high and very high resolution satellite stereo imagery. Dissertation: Universität Osnabrück.

Tian J., Reinartz P., d'Angelo P., Ehlers M. (2013): Region-based automatic building and forest change detection on Cartosat-1 stereo imagery. In: ISPRS Journal of Photogrammetry and Remote Sensing 79 (2013), S. 226-239.

Tian J., Reinartz P. (2013): Fusion of multi-spectral bands and DSM from WorldView-2 Stereo imagery for building extraction. Proceedings of the JURSE 2013, April 21-23, 2013 – Sao Paulo – Brazil. elib.dlr.de/83039/1/43.pdf (10.04.2015)

Tian J., Cui S., Reinartz P. (2014): Building Change Detection Based on Satellite Stereo Imagery and Digital Surface Models. In: IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, Vol. 52, No 1, January 2014, S. 406 – 417.

Tian J., Reinartz P., Dezert J. (2015): Building change detection in satellite stereo imagery based on belief functions. http://www.onera.fr/sites/default/files/u523/C106-2015-JTPRJD-preprint.pdf (04.04.2015)

Ünsalan C., Boyer K.L. (2005): A system to detect houses and residential street networks in multispectral satellite images. In: Computer Vision and Image Understanding 98 (2005), S. 423 – 461.

Vincent L., Soille P. (1991): Watersheds in digital spaces: an efficient algorithm based on immersion simulations. In: IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 13, no. 6, S. 583–598.

Walter V. (2004): Object-based classification of remote sensing data for change detection. In: ISPRS Journal of Photogrammetry and Remote Sensing, vol. 58, no. 3-4, S. 225-238.

Zhou W., Troy A. (2008): An object-oriented approach for analysing and characterizing urban landscape at the parcel level. In: International Journal of Remote Sensing. Vol. 29, No. 11, 10 June 2008, S. 3119-3135.

Eidesstattliche Erklärung

Ich versichere, dass ich die vorliegende Arbeit ohne fremde Hilfe und ohne Benutzung anderer als der angegebenen Quellen angefertigt habe, und dass die Arbeit in gleicher oder ähnlicher Form noch keiner anderen Prüfungsbehörde vorgelegen hat. Alle Ausführungen der Arbeit, die wörtlich oder sinngemäß übernommen wurden, sind als solche gekennzeichnet.

Metzlaff, Lukas

Augsburg, 09.05.2015