PERSISTENT SCATTERER AIDED FACADE LATTICE EXTRACTION IN SINGLE AIRBORNE OPTICAL OBLIQUE IMAGES

L. Schack[a,∗], U. Soergel[b], and C. Heipke[a]

[a] Institute of Photogrammetry and GeoInformation, Leibniz Universität Hannover, Germany - {schack , heipke}@ipi.uni-hannover.de
[b] Institute of Geodesy, Technische Universität Darmstadt, Germany - soergel@geod.tu-darmstadt.de

WG VII/2

KEY WORDS: Urban, Fusion, SAR, Imagery, Extraction, Matching, Pattern

ABSTRACT:

We present a new method to extract patterns of regular facade structures from single optical oblique images. To overcome the missing three-dimensional information we incorporate structural information derived from Persistent Scatter (PS) point cloud data into our method. Single oblique images and PS point clouds have never been combined before and offer promising insights into the compatibility of remotely sensed data of different kinds. Even though the appearance of facades is significantly different, many characteristics of the prominent patterns can be seen in both types of data and can be transferred across the sensor domains. To justify the extraction based on regular facade patterns we show that regular facades appear rather often in typical airborne oblique imagery of urban scenes. The extraction of regular patterns is based on well established tools like cross correlation and is extended by incorporating a module for estimating a window lattice model using a genetic algorithm. Among others the results of our approach can be used to derive a deeper understanding of the emergence of Persistent Scatterers and their fusion with optical imagery. To demonstrate the applicability of the approach we present a concept for data fusion aiming at facade lattices extraction in PS and optical data.

1. INTRODUCTION

1.1 Motivation

In recent years Persistent Scatterer Interferometry (PSI) has become a standard technique for subsidence monitoring of urban areas. Movements at the scale of few millimeters per year can be measured in the sensor line of sight. Nevertheless, to conduct a precise monitoring of defined building structures the reflection process leading to Persistent Scatterers (PS) has to be understood in great detail. Only in this way the correspondence between the PS data point and the equivalent structural element of the building can be ensured. Optical oblique imagery lends itself as a valuable source of information for this process due to a similar sensing direction. In this paper we present a method to establish a link between both types of data by exploiting regular structures at facades. We formulate this process as the extraction of facade regularities in single oblique aerial images with the help of PS grouped to lattice objects.

The establishment of a direct link between Synthetic Aperture Radar (SAR) data and optical imagery in urban areas bears a variety of benefits. On the one hand, one can gain insights into the physical nature of Persistent Scatterers and a better understanding what kinds of object geometries induce them in man-made areas. On the other hand, to establish a connection between both kinds of data offers also benefits from inherent advantages of the individual sensor principles and spectral domains. While a SAR is nearly weather independent, optical data feature a straightforward interpretability by human operators as well as a higher geometrical resolution. The combined use offers also more practical applications, such as change detection in a set-up when a post-event SAR acquisition shall be compared and referenced to a pre-event optical oblique stereo pair, given that only parts are changed. Although these applications are not developed yet for the combination of SAR and oblique imagery and are out of the scope of this paper, the presented work constitutes one important step towards their realization.

1.2 Related work

Our method combines two recent research topics: facade extraction and data fusion. The first is well established and based on a variety of data sources. We treat data fusion only in the context of SAR and optical data. To put our contribution into context we cite some examples of current developments in both fields.

Together with the commercial emergence of oblique camera systems different approaches of deriving building models from this data source were presented. (Xiao et al., 2012) uses the combination of line features like window edges or eaves and height information derived by dense image matching. The approach is based on a fundamental assumption about the typical orthogonal alignment of windows at a facade: The most dominant edges are assumed to belong to vertical and horizontal edges of windows or the facade itself. The results of different viewing aspects are then combined as building hypotheses in a joint framework. To simplify this step, the final facade representation is reduced to the level of planes. A combination of optical close range images and terrestrial laser scanner point clouds is presented in (Pu, 2010). Here, a data base of typical facade models is created first and then applied to new data. The close range imagery is then used to render the building model derived from the laser scanning point cloud by extracting and matching line features in both data types.

Building models can also be derived solely from SAR data. Especially, the recent TomoSAR (Zhu and Bamler, 2010) approach yields high density point clouds which allows for applications comparable to those of LiDAR point clouds. A method of fitting planes and other geometrical primitives in those point clouds is described in (Shahzad and Zhu, 2015). The point cloud is segmented into planes and curved faces without explicitly modeling any specific facade structure.

∗Corresponding author.

(a) Section of an aerial image representing build-up class city center.

(b) Section of an aerial image representing build-up class mixed area.

Figure 1. Examples of the two characteristic building area types.

A set of rules formulated as a grammar is used to describe terrestrial laser point clouds and optical images in (Ripperda and Brenner, 2009) and (Becker, 2009). The explicit formulation of facade structure allow for the description of single windows, doors, and repetitions thereof, but needs a huge amount of training data to define the grammar in the first place.

There are many approaches to fuse SAR data with optical imagery. The joint data analysis is mostly done using time series to classify land cover, e.g. (Chureesampant and Susaki, 2012), (Xu et al., 2010). (Wegner et al., 2008) uses line features in SAR and optical nadir images to establish the correspondence. First investigations regarding the regularity of Persistent Scatterers at building facades are presented in (Schunert and Soergel, 2012). Instead of lattices only horizontal lines of PS were considered. (Gernhardt et al., 2014) incorporates three dimensional building models to evaluate the localization accuracy of PS at facades.

In this paper we present for the first time a method to extract regular facade structures using single oblique images and three-dimensional information derived by grouping PS to lattices corresponding to single facades. The focus lies on the supportive use of SAR data for this task and is one important step towards a deeper understanding of the physical nature of Persistent Scatterers.

2. METHOD

2.1 Urban regularity in optical oblique images

Our approach is based on the assumption that facades in urban areas are characterized by regular and repetitive patterns of facade elements like windows or balconies. This expectation appears trivial due to everyday experience. However, to systematically exploit the present regularity we first need to analyze their appearance in the oblique imagery. In order to do so, representative

scenes of the used investigation area (city center of Berlin) are labeled manually. Images were selected to cover two types of built-up areas. More precisely, we distinguish between these two classes:

- City center, characterized as densely built-up areas of predominantly multistory buildings.

- Mixed areas, featuring a mixture of multistory buildings, town houses of typically 3 to 6 stories as well as single detached houses.

In the city area of Berlin both types of built-up areas can be found. Figure 1 shows examples for both built-up area classes. The facades were classified as showing a regularity if a lattice of at least 3×3 elements was present. Furthermore, the patterns were only considered to be regular if they were not interrupted by other structures like a row of balconies inside a lattice of windows, because our approach is not capable of capturing such complex patterns. The evaluation comprised the ratio of image area to the whole image, the ratio of facades with regularity to the total amount of facades visible in the data, and the median of facade sizes in terms of repeating elements, i.e. the number of single windows at one facade. Our intention is to underline that due to modern architectural style preferring perpendicularity, regularity, and symmetry in large cities like Berlin usually a plethora of periodic facade structures can be found.

Type	facade area [%]	regular facades [%]	# Median regularity
city center	19	32	35
mixed	20	16	20

Table 1. Overview of some key figures describing the occurrence of regular facades in the airborne oblique imagery under investigation. The second column lists the percentage of image area which is covered by a facade, independently whether regularity is present or not. The proportion of facades with a regularity is given in the third column followed by the median of the number of facade elements.

For the evaluated airborne oblique imagery of Berlin a significant percentage of regular facades can be found in the data (see Table 1). In the city center nearly one third of facades shows regular patterns. Even though, the percentage of facade areas compared to the whole image is similar in both types of city environments, the proportion of regular facades is much higher in the city center. This is due to the fact that most high-rise multistory buildings show a very regular facade structure compared to smaller town houses with more individual and interrupted facade patterns. This is also reflected in the smaller number of facade elements (fourth column). Our approach which depends on regular structures seems therefore more suitable for centers of big cities. Note that the overall area of regular facades compared to the whole image is rather small with ~ 3 to 7%. This means that extracting windows from the analyzed oblique images is a very challenging task. Our method uses present PS lattices to narrow down the search space drastically to areas of interest where facade structures can likely be found.

2.2 Facade extraction in Persistent Scatterer point clouds

Given a Persistent Scatter point cloud of an urban area we aim at describing as many PS as possible as part of a facade lattice. The workflow for achieving this goal is as follows (see (Schack and
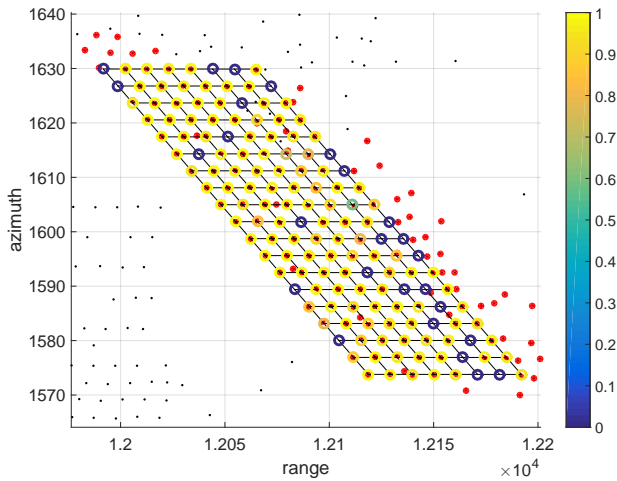
Figure 2. Exemplary result of lattice fitting in SAR data. The color of the lattice nodes encodes the reliability measure, which expresses how well a lattice point describes the closest data point, i.e. a Persistent Scatterer. All PS belonging to the facade are marked in red. Black points are PS of different facades. Some lattice nodes with values close to zero are discernible.

Soergel, 2014) for details): First, the PS point cloud is segmented into clusters of points assumed to belong to the same facade. The segmentation utilizes the normal of the local plane through the 10 nearest neighbors at every PS to distinguish different facade orientations as well as the mutual Euclidean distance to separate facades of the same orientation. Second, a lattice is fitted into the point sets. The parameters to estimate are the two lattice spanning vectors (two directions and spacings), the extent of the lattice, and a measure for the quality of the lattice representation. The resulting lattices are given in two-dimensional SAR as well as 3D object coordinates. Therefore, many properties regarding the geometry of the facade can be derived and further exploited in the optical image domain. Among these are the facade normal, the lattice spanning vectors defining the direction and distance of horizontally (denoted as t_1) and vertically (denoted as t_2) aligned facade objects (mainly window corners), and the outlines of all facades. All these geometrical features are used to make the lattice extraction in single optical oblique images easier. Figure 2 shows a result of the lattice fitting process in SAR data. Several holes in the lattice can be identified. Reasons for missing PS are often occlusions, temporal decorrelation like an opened window during some acquisitions of the PSI time series, or interfering scatterers in the same resolution cell. In section 3.4 we present an approach how a fused lattice can handle such holes.

2.3 Facade extraction in oblique optical imagery

2.3.1 Preprocessing: segmentation and rectification
To simplify the lattice extraction in optical oblique imagery we preprocess the image data in two steps. First and similar to the PS point cloud processing, we divide the oblique image into segments belonging to single facades. Second, image segments are rectified in a way that the horizontal and vertical alignment correspond to the x- and y-image coordinate axes, respectively. Both preprocessing steps are supported by the SAR data.

The separation into segments belonging to single facades is performed by projecting the buffered facade outline (found in the PS data) into the oblique image using the given orientation parameters. This step is crucial and distinguishes the task of facade detection (finding the facade in the whole image) from facade

extraction (describing the facade in the image given the approximate position). As presented in section 2.1 only a very low percentage of the airborne image is covered by the regular facade. With the help of available PS lattices the areas of presumed facades are known. Therefore, available SAR data can help to circumvent the detection task.

To rectify the facade image, we use the same assumptions as in (Xiao et al., 2012), namely we expect high contrast edges along horizontal and vertical facade structures like windows or balconies. Thus, we compute the edge image using the well established Canny operator (Canny, 1986). The t_1 and t_2 vectors are projected into the edge image given the absolute orientation of the airborne oblique image. Since the SAR imaging geometry is different compared to the projective image mapping and also some processing steps are conducted which could add geometrical errors, it is crucial to determine the exact direction of the horizontal and vertical alignment of facade elements in the oblique image. In order to do so, we perform the Hough transformation around the t_1 and t_2 directions separately to obtain precise directions in image coordinates. The rectification is then performed via a two-dimensional affine transformation which maps the three points $\{(0,0); t_1; t_2\}$ onto $\{(0,0); (1,0); (0,1)\}$. Figure 3c shows an example of a segmented and rectified facade image.

2.3.2 Extracting regularity via cross-correlation
The fundamental assumption of our approach is that a sufficient number of facades exist which consist of repetitive elements of similar appearance in the oblique image. Therefore, we aim at dividing the rectified facade into equally sized parts along the x- and y-axes. To capture this kind of regularity the concept of cross correlation is widely used and also offers a direct measure of similarity. Again, the additional SAR data play an important role. The size of the initial template I_0 is defined by the two spanning vectors of the SAR lattice (t_1 and t_2) multiplied by a factor larger than one to ensure that the template is at least as wide as the repeating facade element. The template is taken from the center of the rectified image.

To produce a template that allows for a robust cross-correlation we use the following iterative procedure: The whole image is cross correlated with I_0 by sequentially shifting the template over the image and computing the correlation coefficient for every pixel. The local maxima along the x- and y-axes are selected if they exceed a certain threshold T and the spacings in between are integrated into an accumulator. The surrounding of every selected maximum is added to the template I_0. This is performed until no valid peak can be found or if a maximum number of iterations is reached. The most often voted spacing is then taken as the spacing in the image domain in x- and y-direction, respectively. The template I_0 (comprising the sum of all neighborhoods at local maxima) is normalized to obtain a mean template I_m describing an average of the model of the repeated facade element. Figure 3 shows this procedure for one example. The sensitivity of our approach regarding the parameter T is investigated in section 3.2.

Having divided the facade into repeating segments of the apparently same structure, the extent of the regularity has to be estimated in terms of integer repetitions of the spanning vectors in a second step. This step is necessary because the above procedure (finding local maxima of cross correlation) does not guarantee complete lattices, i.e. rows or columns of regular facade structures can be missing. In order to estimate the correct lattice extent, we use the computed mean patch (exemplarily shown in Figure 3b) and work only on facade segments. We formulate this estimation as a binary classification or figure-ground separation.
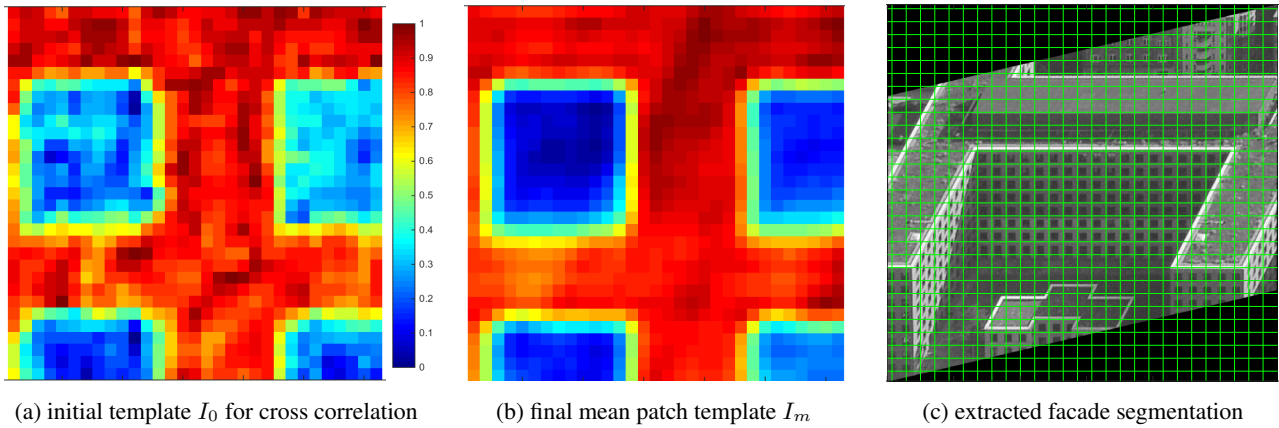
(a) initial template I_0 for cross correlation   (b) final mean patch template I_m (c) extracted facade segmentation

Figure 3. Process of extracting the facade regularity via cross correlation. **(a)** shows the initial template which is derived from the SAR lattice. The color codes the normalized gray values. **(b)** depicts the mean patch after the iterative procedure described in the text. A higher contrast compared to the initial patch is discernible. **(c)** illustrates the extracted lattice regularity. Note that the extent of the regularity is not known at this time and, therefore, is continued to the image borders.

Assuming that the facade is composed of connected regularity segments we consider the facade to be a simply connected set of segments in the middle of the image while all other segments do not belong to the facade. Again, the correlation coefficient of the segments with the mean patch measures the magnitude of similarity. To deduce a threshold which separates the two classes we follow the method of (Otsu, 1979) which minimizes inner-class variances. To obtain a simply connected subset of components we apply morphological operators on the thresholded result. Figure 4 shows an example for this processing chain.

2.3.3 Window lattice model A further exploitation of the regularity at facades is to estimate a window lattice in the image. The benefit of this step can be motivated in two ways: First, the extraction of regularity based on cross correlation can be stabilized, assuming that the alignment of windows at the facade is also regular.

Second, for many applications knowledge about the appearance of windows at the facade constitutes very valuable information. For instance, most visualization applications require texture information. Beyond that, explicit window models offer geometric information which can be further developed in the context of scene interpretation. Also, the above motivated fusion of PS with optical imagery can be simplified. It is well understood that Persistent Scatterers are induced, among others, by trihedral corners with a minimal side length of about 8 cm for recent SAR sensors like TerraSAR-X (Bamler et al., 2009). The window sill, frame, and adjacent wall often form such a trihedral object. Therefore, explicitly modeling window corners is equivalent to creating high-potential candidates for building structures inducing Persistent Scatterers. To establish a link between PS and their correspondences in optical imagery is thus significantly alleviated if the search space of possible matches is reduced to some corners.

Figure 5 shows a schematic view of the window model. The whole lattice for one facade is illustrated in (a). The origin of the whole lattice is marked by a red point P_0. Bounded integer multiples of the spacing Δx and Δy determine all lattice positions (green points). At every lattice position a window is situated with width δx and height δy. We restrict our window lattice model by the following constraints:

- The horizontal as well as vertical spacing are constant for one facade. This means that the distance between neighboring windows is constant (Δx and Δy in Figure 5a are

constant per facade). This excludes curved facade surfaces as well as irregular or more complex patterns. These issues are subject to future studies.

- Every window of a lattice has the same height and width (δx and δy in Figure 5 are constant per lattice).

- The lattice has to be complete in the sense that no complete column or row is allowed to be missing inside the lattice. Note that single windows are allowed to be absent in the lattice representation.

Given these restrictions, a window lattice can be determined by six parameters: lattice spacings Δx and Δy, window width δx and height δy, and two-dimensional lattice origin P_0.

Since window edges often coincide with a strong gradient magnitude compared to their surroundings in optical images (e.g., due to different reflection properties of glass and wall materials), we use the edge image from the processing step above to derive an optimal window lattice according to the outlined restrictions. The quality of the fit is measured via the amount of pixels described by the window lattice model which coincide with edge pixels. The set of all pixels described by the window model in Figure 5b are

$$
\begin{aligned}
W(t_0, \delta_x, \delta_y) = \{ p \,|\, & (p_x = t_{0x} \land t_{0y} \le p_y \le t_{0y} + \delta_y) \\
& \lor (p_x = t_{0x} + \delta_x \land t_{0y} \le p_y \le t_{0y} + \delta_y) \\
& \lor (t_{0x} \le px \le t_{0x} + \delta_x \land p_y = t_{0y}) \\
& \lor (t_{0x} \le px \le t_{0x} + \delta_x \land p_y = t_{0y} + \delta_y) \}
\end{aligned}
\tag{1}
$$

where the color of the single terms correspond to the equivalent parts of Figure 5b. $W(t_0, \delta_x, \delta_y)$ comprises all pixels defined by the rectangle with origin t_0, width δx, and height δy describing one single window. The set of lattice points L, defining all origins of single windows can be expressed as

$$
\begin{aligned}
L(P_0) = \{ P_0 + a \cdot \Delta x + b \cdot \Delta y \,|\, & a, b \in \mathbb{N}, \\
& a \le a_{max}, \\
& b \le b_{max} \}
\end{aligned}
\tag{2}
$$

where a and b are the integer repetitions (including 0) of the spacings Δx and Δy. The extent of the lattice is bounded by a_{max} and b_{max}.

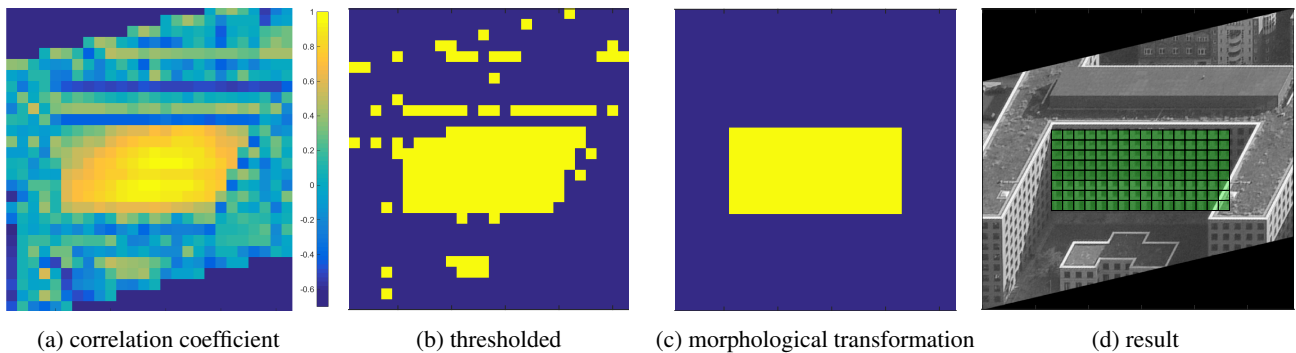| (a) correlation coefficient | (b) thresholded | (c) morphological transformation | (d) result |

Figure 4. Process of estimating the extent of the facade regularity. **(a)** shows the correlation coefficient of every segment. The color codes the correlation coefficient. The facade structure as well as the occlusion by the orthogonal building part is visible. **(b)** displays thresholded binary image. The inter-class variance maximization yielded a threshold of 0.34 for this example. Patches with correlation coefficient about this threshold are marked in yellow. **(c)** depicts the thresholded result after a sequence of opening and closing. **(d)** shows the result of the extent estimation overlayed on the oblique image.

| (a) Schematic view of the facade model. | (b) Schematic view of the window model. |
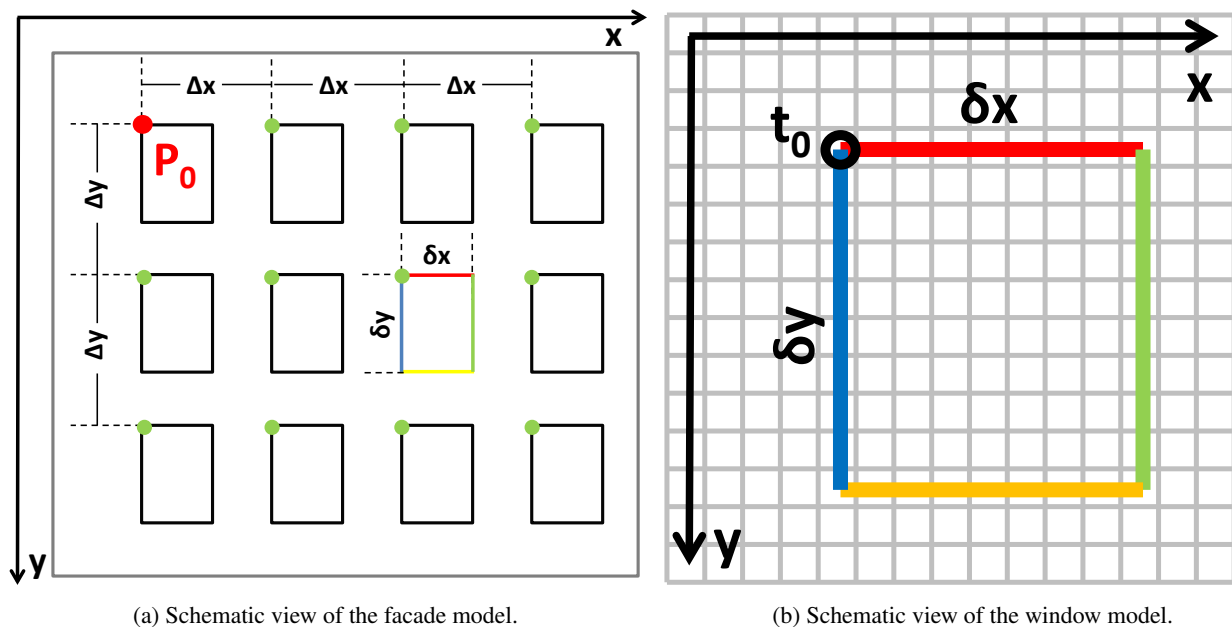
Figure 5. **(a)** The facade is composed of single windows as depicted in (b). All origins of single windows (denoted t_0 in (b)) are marked with green points. The lattice is defined by the two spacings Δx and Δy, and the lattice origin P_0. One window is exemplarily colored as in (a). **(b)** The discrete pixel locations of the window edges are determined by the origin of this specific image t_0 and the two parameters δ_x and δ_y defining the width and height of the window. The color correspond to the terms in equation (1).

We then maximize the amount of pixels described by the window lattice model that coincide with edge pixels:

$$\operatorname*{maximize}_{P_0, \delta_x, \delta_y} \sum_{e \in E} e = \left\{ \begin{array}{ll} 1, & e \in W(L(P_0), \delta_x, \delta_y) \\ 0, & e \notin W(L(P_0), \delta_x, \delta_y) \end{array} \right. \quad (3)$$

where E is the set of all edge pixels e. If the pixel is also covered by the window lattice, it is part of the sum. Note that the spacings Δx and Δy are not part of the optimization procedure since the vertical and horizontal regularity is already known from the previous processing step.

The optimization of (3) puts two requirements on the solving algorithm: First, all parameters are integers since we operate on pixels. Second, if one state of parameters is found, it cannot be inferred where the next better state is located, because due to the binary edge image no meaningful derivative exist. To circumvent this problem either the edge image could be smoothed, e.g. with a Gaussian kernel, to approximate a derivative, or, an optimization approach which is not based on the gradient could be used. Taking into account both aspects, we use a genetic algorithms with standard parameters (population size: 50, crossover rate 0.8, migration rate 0.2, elite fraction: 0.05) to find a globally maximum solution for (3). It is a derivative-free method which can also handle integer problems (Deep et al., 2009). According to the notation of genetic algorithms, the parameters to optimize $(P_0, \delta_x, \delta_y)$ are called the genome while a solution for given parameters is called the fitness function. The key concept of a genetic algorithm is to generate a population of several members with initially random genome and then derive descendents of pairs from the actual generation. Every member of the new generation is scored by the fitness function and only the fittest are considered for building the next generation. To avoid getting caught in local minima a certain percentage of new genome combinations are exposed to mutations. Nevertheless, the used genetic algorithm cannot guarantee global optimality and therefore, the optimization is performed several times in parallel to improve the chance to find a solution close to the global maximum. The parameter set yielding the highest fitness score is then chosen as the resulting window lattice model.

3. EXPERIMENTS

3.1 Data

We use data from the city of Berlin to test the proposed method. The Persistent Scatter point cloud was derived from a stack of 54 TerraSAR-X High-Resolution Spotlight acquisitions. It was processed with the GENESIS PSI processor (Adam et al., 2003) by Stefan Gernhardt from Technische Universität München. The data were manually separated into regions of interest which are building complexes or single city blocks. The airborne imagery consists of oblique images covering an area of approximately $0.6 - 0.7 \text{ km}^2$. The images were captured in autumn 2010 and have a ground sampling distance of approximately 15cm in the scene center, they were provided by BLOM UK.

3.2 Parameter considerations

A crucial parameter of our approach is the value T which governs the decision whether an image segment is similar enough to the mean patch to be included in the set of regular segments. To evaluate the sensitivity of this parameter, we run several experiments on the same building complex discussed in section 3.3.1 with varying values for T, while all other parameters remained constant. As a measure of extraction quality, the number of correctly extracted windows is counted having manually labeled the image as ground truth before. A window is considered as extracted when the manually labeled window center lies inside the
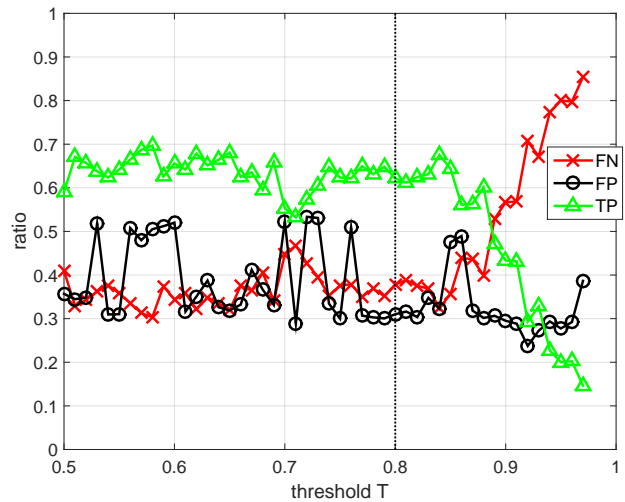

Figure 6. Influence of the threshold T on the quality of the window lattice extraction result. The true positive rate (TP) is plotted in green and measures the ratio of correctly extracted windows compared to the total amount of all windows in the ground truth data. The false positive (FP) rate, black, is the ratio of extracted windows which are not in the ground truth data compared to all extracted windows. The red plotted false negative (FN) rate measures the ratio of how many windows of the ground truth data are missed.

polygon of the window according to the window lattice model of equation (2). We differentiate between correctly extracted windows (true positive, TP), falsely identified windows (false positive, FP), and missed windows (false negative, FN). The threshold T is tested for values between 0.5 and 1 with a step size of 0.01.

Figure 6 shows the three quality measures as a function of T. TP and FN are nearly constant in an area of $0.5 \leq T \leq 0.85$. In this interval the true positive rate varies between 0.6 and 0.7 which means that the majority of windows is extracted correctly. Of course, this number has to be scaled with the false positive number which varies between 0.3 and 0.55. The second part of the graph begins to the right of $T = 0.85$ where the true positive rate rapidly drops due to the fact that a too rigorous threshold prevents the algorithm to create a robust mean patch for the cross correlation process. For the following case studies a threshold of $T = 0.8$ is chosen because the gap between the TP and FP rate is quite big for a certain range around this value.

3.3 Case studies

To show the applicability of the presented approach, three case studies are presented in this section. Some basic characteristics of our approach which are independent on the used data are discussed.

3.3.1 Park Kolonnaden Figure 7 shows an accumulation of preeminently regular facades at the building complex 'Berliner Park Kolonnaden'. Most facades are extracted correctly in terms of the vertical and horizontal window spacings, the window geometry and also the extents of the window lattices. Since our approach enforces complete lattices, some windows are extracted which are not supported by the image data. Such areas are marked with green arrows. Often this is due to occlusions in the image. Nevertheless, the lattice information can be used to infer such positions. The white arrow points to a facade which is further discussed in section 3.4.
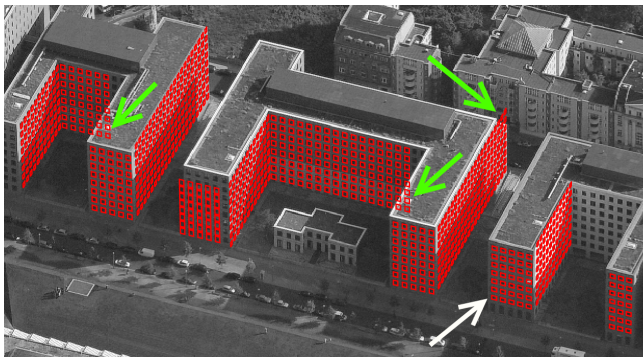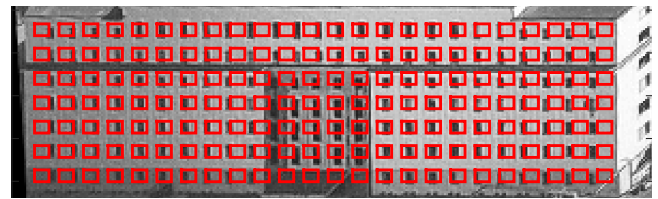
Figure 7. Building complex 'Berliner Park Kolonnaden'.
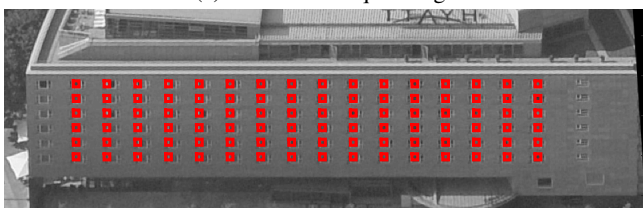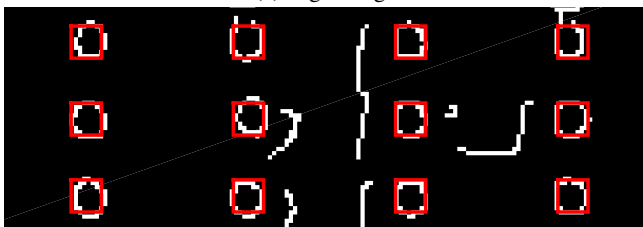

(a) facade in oblique image

(b) facade in rectified oblique image

(c) edge image

(d) detail of edge image

Figure 8. Hotel. **(a)** depicts the facade in the original oblique image. **(b)** shows the rectified facade overlaid with the extraction result in red. **(c)** displays the corresponding edge image. **(d)** is a zoom in of (c) for a better understanding.

(a) rectified oblique image

(b) spacing

Figure 9. **(a)** shows the rectified building facade overlayed with the extraction result of the windows lattice model. **(b)** shows the result of the cross correlation step explained in section 2.3.2.

3.3.2 Hotel The influence of the window structure in terms of the resulting edge image can be understood with the help of the example in Figure 8. Windows at this facade are tripartite, separated by a vertical frame inside the overall window. The inner frame features a stronger gradient magnitude resulting in apparently too small windows, as the extraction results depict. The detailed view in Figure 8d shows that the window fitting is correct, given the edge image.

3.3.3 Axel-Springer-Strasse The extent of the regularity at the facade shown in Figure 9 can be estimated correctly while the window geometry is wrong. One reason is the incorrectly estimated spacing, especially in horizontal direction. Figure 9b depicts a wandering of the windows relative to the green spacing lines. This leads to a violation of the assumptions of the used window lattice model described by equation (2) as the origin of every window (t_0 in Figure 5) varies within the facade patch defined by the spacing.

3.4 A fusion outline

To show the applicability of the presented approach we present a rough outline of a fusion scheme in this section. To make the lattice node quality comparable across the sensor domains we normalize the reliability measure, described in (Schack and Soergel, 2014) as well as the correlation coefficient between each lattice tile and the mean patch (see section 2.3.2). Note, that this fusion outline does not allow for any statistical statements, but is supposed to give a general idea of how the sensor fusion can be performed.

An image of the facade under investigation is given in Figure 10. The fusing result is shown in Figure 11. The PS lattice in (a) exhibits three nodes which are not supported by the data resulting in a low value for the reliability measure. All other nodes have values close to one and, thus, can be assumed to reliably describe the SAR data. The optical lattice has only one node with a low correlation coefficient resulting from a reflection effect (visible in Figure 10). All other nodes are in the range between 0.6...1 and are assumed to be reliable as well. From the joint interpretation of both sensor domains, and assuming that the result of one sensor is enough evidence, a more complete lattice than using only one data source can be inferred. Since there is no evidence in the optical data which explains the absence of those three points with low reliability measure (color coded black in Figure 11a)
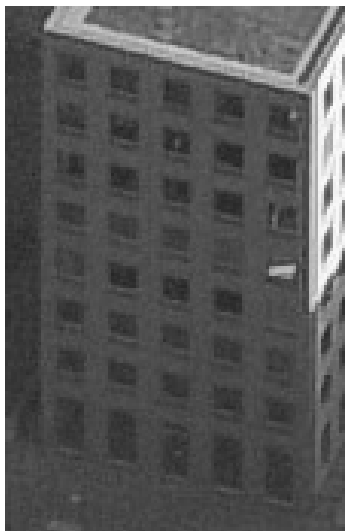
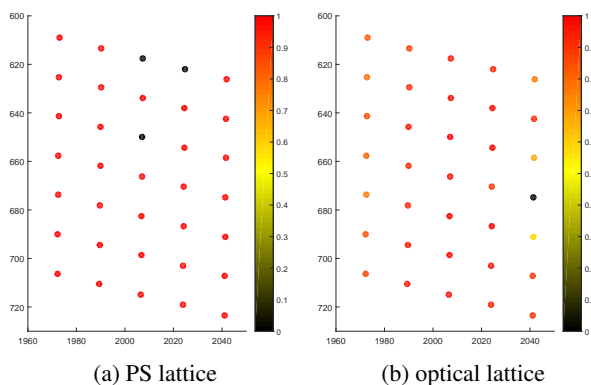Figure 10. Facade used for fusion example. In the right most window column a reflection is visible.

(a) PS lattice   (b) optical lattice

Figure 11. Corresponding points of the facade marked with a white arrow in Figure 7. Figure 10 is the corresponding close-up image. **(a)** are the points color-coded according to the reliability measure described in (Schack and Soergel, 2014). **(b)** shows the corresponding points but color-coded according to the correlation coefficient described in section 2.3.2.

in the SAR data, it could be helpful to further investigate the PS processing at these locations.

4. CONCLUSION AND OUTLOOK

This paper describes a method for facade extraction from single airborne oblique imagery with the aid of Persistent Scatter point cloud data. The facade extraction is based on the observation that many building faces show regular patterns of windows, balconies, or other structures. This regularity is exploited to establish a link between remotely sensed data of these very different sensors. Three case studies show the applicability of the approach. The presented method is not meant to compete with state-of-the art facade extraction algorithms for stereo oblique or LiDAR data but gives some interesting insights for the challenging task of sensor fusion in urban areas. Explicitly modeling a window lattice reduces the possible candidates of pixels corresponding to distinct PS significantly.

In future work we plan to use the presented method to systemati-

cally investigate facades and their correspondences in both types of data. By doing so, we hope to better clarify the behavior of PS at facades and their actual presence or absence in terms of nodes of the representing lattices. Furthermore, one aim is to extend the presented approach for fusing PS and optical imagery in different ways. As an example, at the moment the lattice extraction is limited to a very restricted form of regularity. To deal with more general cases of facade regularity a more sophisticated method like grammars can be integrated as long as the lattice nodes have the same size. The result can then be evaluated with a measure similar to the correlation coefficient. Another extension is to incorporate stereo image pairs which gives direct three-dimensional information and allows for fusion in object space.

ACKNOWLEDGEMENTS

REFERENCES

Adam, N., Kampes, B., Eineder, M., Worawattanamateekulb, J. and Kircher, M., 2003. The development of a scientific permanent scatterer system. In: Proceedings of ISPRS Workshop High Resolution Mapping from Space 2003, pp. 6–11.

Bamler, R., Eineder, M., Adam, N., Zhu, X. X. and Gernhardt, S., 2009. Interferometric Potential of High Resolution Spaceborne SAR. Photogrammetrie - Fernerkundung - Geoinformation 2009(5), pp. 407–419.

Becker, S., 2009. Generation and application of rules for quality dependent facade reconstruction. ISPRS Journal of Photogrammetry and Remote Sensing 64(6), pp. 640–653.

Canny, J., 1986. A computational approach to edge detection. IEEE Transactions on Pattern Analysis and Machine Intelligence 8(6), pp. 679–98.

Chureesampant, K. and Susaki, J., 2012. Land cover classification using multi-temporal SAR data and optical data fusion with adaptive training sample selection. In: 2012 IEEE International Geoscience and Remote Sensing Symposium, Ieee, pp. 6177–6180.

Deep, K., Singh, K. P., Kansal, M. and Mohan, C., 2009. A real coded genetic algorithm for solving integer and mixed integer optimization problems. Applied Mathematics and Computation 212(2), pp. 505–518.

Gernhardt, S., Auer, S. and Eder, K., 2014. Persistent scatterers at building facades – Evaluation of appearance and localization accuracy. ISPRS Journal of Photogrammetry and Remote Sensing 100, pp. 92–105.

Otsu, N., 1979. A threshold selection method from gray-level histograms. IEEE Transactions on Systems, Man and Cybernetics 9(1), pp. 62–66.

Pu, S., 2010. Knowledge based building facade reconstruction from laser point clouds and images. PhD thesis, University of Twente.

Ripperda, N. and Brenner, C., 2009. Application of a formal grammar to facade reconstruction in semiautomatic and automatic environments. In: Proceedings of the 12th AGILE International Conference on Geographic Information Science 2009, Hannover, pp. 1–12.

Schack, L. and Soergel, U., 2014. Exploiting Regular Patterns to Group Persistent Scatterers in Urban Areas. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 7(10), pp. 4177–4183.

Schunert, A. and Soergel, U., 2012. Grouping of Persistent Scatterers in high-resolution SAR data of urban scenes. ISPRS Journal of Photogrammetry and Remote Sensing 73, pp. 80–88.

Shahzad, M. and Zhu, X. X., 2015. Robust Reconstruction of Building Facades for Large Areas Using Spaceborne TomoSAR Point Clouds. IEEE Transactions on Geoscience and Remote Sensing 53(2), pp. 752–769.

Wegner, J., Inglada, J. and Tison, C., 2008. Automatic Fusion of SAR and Optical Imagery based on Line Features. In: European Conference on Synthetic Aperture Radar, pp. 171–174.

Xiao, J., Gerke, M. and Vosselman, G., 2012. Building extraction from oblique airborne imagery based on robust façade detection. ISPRS Journal of Photogrammetry and Remote Sensing 68, pp. 56–68.

Xu, M., Xia, Z., Zhang, F., Li, K. and Xie, C., 2010. Multi-Temporal Polarimetric SAR and Optical Data Fusion for Land Cover Mapping in Southwest China. In: International Conference on Multimedia Technology (ICMT) 2010, Ningbo, pp. 1–4.

Zhu, X. X. and Bamler, R., 2010. Very high resolution spaceborne SAR tomography in urban environment. IEEE Journal of Geoscience and Remote Sensing 48(12), pp. 4296–4308.