PFG 2013 / 5, 0511 – 0521
Stuttgart, October 2013

Article

Cornice Detection Using Façade Image and Point Cloud

Wolfgang Brandenburger, Euskirchen, Martin Drauschke & Helmut Mayer, Neubiberg

Keywords: architectural model, LOD3, façade interpretation, window detection, RANSAC

Summary: We propose an approach for the automatic detection of cornice structures above windows. It combines the edges from a rectified façade image and the geometrical information of a 3D point cloud derived from a set of images, e.g. using Mayer et al. (2012). Based on the detection of the façade plane in the point cloud and the segmentation of its outline we detect windows to restrict the search space for cornices. Analysing the 3D points as well as the detected edges we are able to reliably distinguish between triangular, arch-shaped and horizontal cornices. This is demonstrated by experiments for façades from Freising, Munich and Schwabach.

Zusammenfassung: *Detektion von Fensterverdachungen in Fassadenbildern und Punktwolken.* In dieser Arbeit wird ein Ansatz für die automatische Detektion von Fensterverdachungen und Gesims vorgestellt. Dieser kombiniert die Kanten eines entzerrten Fassadenbildes mit der geometrischen Information einer 3D-Punktwolke, die aus einer Bildmenge abgeleitet wird, z. B. nach Mayer et al. (2012). Ausgehend von der Detektion der Fassadenwand als Ebene und der Bestimmung ihrer Begrenzung werden Fenster detektiert, um den Suchraum für Verdachungen einzuschränken. Durch Analyse der 3D-Punkte und der detektierten Kanten kann zuverlässig zwischen bogenförmigen Fensterverdachungen, Dreiecksgiebeln und Gesims unterschieden werden. Das wird in Experimenten auf Fassaden aus Freising, München und Schwabach gezeigt.

1 Introduction

Three-dimensional (3D) city models are often used for tourism or for architecture and city planning. Especially in these domains, building models with details such as façade ornaments can be very important for a realistic visual appearance.

As state-of-the-art, most buildings in larger cities are modelled in a simple way, often with level of detail (LOD) 1 or 2 according to the CityGML standard (Gröger et al. 2012). This means that they only consist of planar walls possibly texturized using a façade image, and simple roof structures. Often only landmarks and important tourist locations of a city are manually modelled with more details as the automatic enrichment of building models with façade details (LOD3) is a very challenging research goal in photogrammetry and computer vision. The Internet of things is another application for high resolution interpreted 3D scenes. According to Leberl et al. (2012), information down to 3 cm resolution is needed for virtual navigation through urban spaces. Therefore, also the recognition of small details on building façades will be required for realistic models of the environment.

Leberl et al. (2012) derive large and highly detailed 3D models of urban spaces from various data sources. Aerial images are used to generate coarse building and landscape models, whereas laser scans and terrestrial images from mobile sensing platforms are employed for the determination of models with more detail. If these results can be merged with future results of automated approaches for the interpretation of façades such as Hohmann et al. (2009), the combination will make automatic

 www.schweizerbart.de
DOI: 10.1127/1432-8364/2013/0196 1432-8364/13/0196 $ 3.00

building modelling feasible both at city scale and at an individual building scale.

Due to the reliance on laser scans, a large budget is required for city modelling using the methods described in Hohmann et al. (2009). In contrast, Mayer et al. (2012) show a more affordable way for obtaining information about 3D geometry. They combine terrestrial images and possibly images from unmanned aircraft systems to generate highly precise point clouds whose point densities are high enough to reconstruct building surfaces (Nguatem et al. 2012).

In this article, we present an approach for detecting cornices at façades using terrestrial images and point clouds derived from such images. It is structured as follows: In section 2, we review recent work on detecting façade parts in images or point clouds. Façade plane and window detection are summarized in section 3, before we present our new approach for cornice detection in section 4. After presenting experiments in section 5, we conclude with suggestions for future work in section 6.

2 Related Work on Automatic Detection of Architectural Elements

In this section, we give an overview on recent work regarding the detection of façade parts in images. Some authors propose to increase the detection rate by making use of the multiple occurrences of many façade elements such as windows. This leads to grammar-based approaches which are mainly used for interpreting entire façades. Because this is a recent trend, we also review work on façade parts in 3D.

Lee & Nevatia (2004) is one of the more recent publications on the automatic detection of windows in images. Their approach is limited to façades with a regular placement of windows, because they analyse cumulative image gradients in column and row direction. Furthermore, their approach is usually applied to low resolution projections of aerial images on façade planes, whose vertical alignment is known. Thus, terrestrial façade images are often too challenging for this approach. Reznik & Mayer (2008) particularly focus on this later type of images. An implicit shape model (ISM) is used to derive window centres and window borders from already recognized window corners. Furthermore, the information criterion from Akaike (1973) is applied for improving the alignment of windows by arranging them in a row or column structure. This approach has two important limitations: Firstly, it does not perform well if windows are different in size, and secondly, windows with round arches cannot be detected. Both approaches, Lee & Nevatia (2004) as well as Reznik & Mayer (2008), analyse individual windows, and the results are integrated into 3D building models including the depth of the windows.

The following three papers only use single façade views without any 3D information. Wenzel & Förstner (2012) propose a bag-of-words approach to detect façade parts. They aggregate segmented edges to curves and compare those with previously learnt aggregates. Thus, façade parts limited to windows, entrances and balconies are identified. Another general approach for detecting visibly distinguishable elements of façades has been proposed by Jahangiri & Petrou (2008). Various blob-like image structures are detected. Autocorrelation between the detected elements is performed to determine candidate vectors for repetitions. Since windows are often the most frequent façade parts, the main modes of the autocorrelation function correspond to repetitive window patterns. Therefore, their regular structure can be used to improve the extraction results with respect to the window alignments. We think that it is possible to detect cornices in high resolution images employing this procedure, but then also many other blob-like structures will probably be obtained as false detections. A similar procedure for detecting repetitive façade parts has been proposed by Zeng et al. (2011). As they focus on larger entities such as windows surrounded by shutters, cornices as well as the plain wall, their approach is not applicable to detect cornices only.

We know about only one work which explicitly tried to detect cornices in images without using 3D information. In an internal note, Šochman (2007) confesses that his workflow using Haar-like features does not work well for detecting triangular cornices with a boosting classifier.

Simon et al. (2011) proposed a grammar-based approach for segmenting a rectified façade image into non-overlapping parts such as windows, balustrades of balconies and doors. The façade is successively decomposed into parts using a grammar specifically designed for Haussmannian buildings in Paris, i.e., buildings constructed in the late 19th century Paris. For modern buildings or buildings of other styles a different grammar must be learnt. Thus, this approach cannot easily be transferred to buildings in other cities and from other epochs. Ripperda (2008) presented a more general approach for façade interpretation, and furthermore, she combines image and depth information from laser scans. She focuses on windows and doors, although her grammar seems to have potential to model the window's surrounding including artistic façade parts such as cornices, analogously to Wonka et al. (2003).

Becker & Haala (2008) employ laser scans to automatically detect windows with mullions and transoms. They decompose the façade into cells and try to find homogeneous parts representing the plain wall, the window pane or parts along 3D edges representing window borders. The cells are merged to model the façade with windows and doors. The depth of these façade parts is determined by reconstructing 3D edges from an image pair. Although Becker & Haala (2008) were able to reconstruct details of windows behind the wall, they did not model anything in front of the wall.

Hohmann et al. (2009) present the project CityFit, which has the goal of an automatic reconstruction of the entire city of Graz from mobile laser scans and images. Although the scope of the project also comprises the modelling of details such as cornices and other ornaments, only results for the automatic segmentation of modern, plain façades have been published so far.

Becker & Haala (2008) and Reznik & Mayer (2008) demonstrate that 3D information is very helpful to reconstruct building façades showing windows slightly behind the wall. Thus, we think it could be also useful to detect façade parts such as cornices, which are outside the main plane corresponding to the wall, particularly in combination with image analysis.

3 Façade Plane and Window Detection

Our approach for façade interpretation is based on images acquired using hand-held consumer cameras. By means of the reconstruction technique of Mayer et al. (2012) we obtain a very precise 3D point cloud with a back projection error in the range of 0.1 to 0.2 pixels. This approach allows the reconstruction of a medium-density point cloud from wide baseline image sets combining feature-based least squares matching and robust bundle adjustment. We obtain only about 1% of the number of 3D points that could be acquired by dense matching, e.g., Hirschmüller (2008). As will be shown in our experiments, this is often enough for achieving good results for detecting cornices.

As a first step for preprocessing the data, one of the images showing a façade of one building is selected and rectified by projecting it to the façade plane (Fig. 1). An individual façade is only interpreted by the related rectified image. In downtown scenes usually only one façade or at most two façades are visible, the latter if the building is situated at a corner

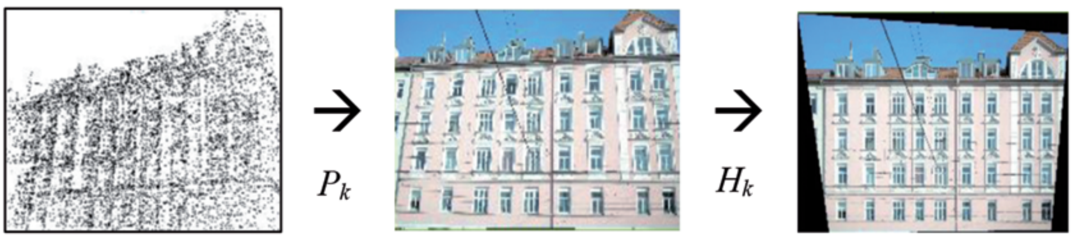

Fig. 1: Projection and homography matrix for mapping 3D points and pixels of the rectified image.

of a block. In the corner case, the façades must be selected one by one, and the preprocessing steps and the cornice detection have to be performed twice. So far, we manually determine a homography to rectify the selected view, but this step can also be done automatically following (Mayer 2007), if the vertical direction of the scene is known. The correspondence between point p_m of the point cloud and the pixel q_n in the rectified image is given by

$$q_n = H_k \cdot P_k \cdot p_m, \tag{1}$$

where P_k is the projection matrix for transferring a 3D point into the kth image, and H_k is the homography for rectifying the kth image (Fig. 1), i.e. relating the image coordinate reference system (crs) to another crs with its (X,Y) plane identical to the façade itself.

Our approach for cornice detection is based on two further automated preprocessing steps, for which the correspondence between the point cloud and the rectified image is assumed to be already established. We start with detecting the façade plane and continue with the detection of windows employing the gradient-based approach of Lee & Nevatia (2004).

We approximately locate the façade plane in the point cloud using the RANSAC-based robust plane fitting approach by Nguatem et al. (2012). RANSAC (Fischler & Bolles 1981) yields correct approximations. However, they are not necessarily of high quality. As a consequence, we improve the plane detection by the following steps, where we also consider the visual appearance of the 3D points near the façade.

Assuming, that we know the scale of the point cloud, we can extract the 3D points which are located at maximum 0.1 m in front of or behind the façade plane. In a following step we filter out all 3D points belonging to small structures such as ornaments and window frames in order to finally define the façade with an accuracy of a few millimetres. This is done by an inspection of the image trying to segment the wall of the façade without ornaments and windows, i.e., we analyse each pixel regarding the fulfilment of three conditions: 1) The intensity values in the neighbourhood have a small variance since unstructured parts of façades are often homogeneous regarding their radiometric intensity. 2) The distance of points in the neighbourhood of the plane is small. 3) The variance of these distances is also small. Additionally, morphological operations are used to remove isolated façade points and smooth the border of the segmented façade (Fig. 2). Using (1) we determine all 3D points lying on the segmented façade's wall and finally complete the plane detection by a least-squares adjustment using all those 3D points leading to a more precise façade plane. As a positive consequence, we are now able to identify 3D points which lie only slightly behind or in front of the plane.

After locating the façade plane, we manually extract the boundary of the façade restricted

Fig. 2: Segmentation of rectified façade image. We show all pixels which correspond to the façade plane in blue. They are homogeneous with respect to the intensity values. The corresponding 3D points are close to the plane and their variance with respect to the distance is low. In magenta we show the results of window detection and in red the detected triangular cornices.

to the upper floors in the rectified image, because windows at ground floor often look differently (RIPPERDA 2008), and we employ the window detection of LEE & NEVATIA (2004). Large façades in city centres and especially of historical buildings usually have regularly arranged windows. Thus, we expect reasonable results when locating windows by adding up gradients in row and column directions. In Fig. 2, the detected windows are shown. Their locations and sizes are not correct regarding the column direction, but the results still lie in an acceptable range. If a façade does not show an array-like window pattern, the gradient-based window detection might be substituted by the corner-based detection at the core of the ISM-based approach (REZNIK & MAYER 2008).

Concluding this section, in the preprocessing steps we have selected the image for façade interpretation, we have rectified it, and we have determined the correspondence between the 3D points of the point cloud and the pixels of the rectified image. Using this information, we are able to exactly detect the façade plane, to extract the upper floors, and to detect the windows.

4 Cornice Detection

4.1 *Object Models and Search Space for Cornices*

Cornices are façade parts which protrude from the façade plane, and lie directly above windows. They have no other function than decoration, supposed to make a façade appear livelier by casting little shadows (KOEPF & BINDING 2005). Typically, the width of a cornice is directly related to the width of the window beneath with the cornice usually being a little bit wider. Furthermore, cornices typically have a moderate height. We consider relative position as well as size for cornice detection.

The three most frequently appearing types of cornices in central European cities are shown in Fig. 3. A *horizontal cornice* consists of a flat and thin roof. Sometimes smaller ornamental structures which protrude from the façade less prominently are below the roof. A *triangular cornice* shows a triangle with a horizontal base and two oblique sides of equal length. Sometimes the base is a little bit behind the upper two roof sides, or it is replaced by artistical ornaments below the upper roof sides. An *arch-shaped cornice* shows the top part of an arch and is nearly always aligned with a horizontal edge below. Sometimes there are ornaments inside the arch. We model each detected cornice by one of the three shapes: a rectangle for horizontal cornices, a triangle for triangularly shaped cornices and a section of a circle above a chord for arch-shaped cornices.

When detecting windows using LEE & NEVATIA (2004) an array structure is obtained. We identify a detected window by the two indices i and j corresponding to the position in the array. Hence, we may parameterize each detected window W_{ij} by the four-tuple (x_{ij}, y_{ij}, w_{ij}, h_{ij}) describing the position of its centre (x_{ij}, y_{ij}) in the image and its spatial extents with width w_{ij} and height h_{ij}. We restrict the search space for cornices to the top border of window W_{ij}, the bottom line of the window above $W_{i-1,j}$ and the two windows $W_{i,j-1}$ and W_{ij+1} to the left and to the right (Fig. 3). If there are no neighbouring windows, the search space is restricted by the boundary of the façade.

We have analysed three strategies for cornice detection which are explained in the following subsections. Using the first strategy, we detect cornices by means of the structure of their appearance in the point cloud, while in

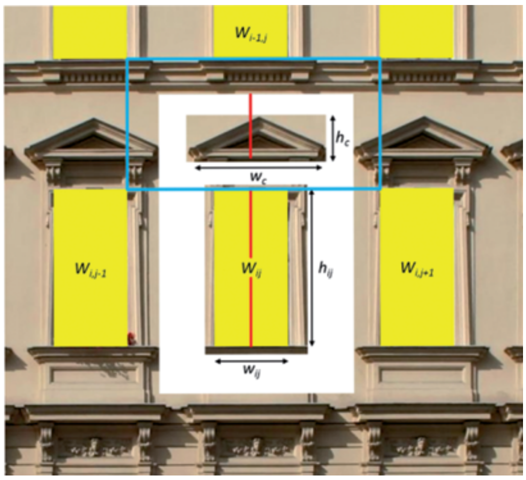

Fig. 3: Clipped façade view with detected windows in yellow and the search space for a cornice above window W_{ij} in blue and the centreline of window W_{ij} in red.

the second we detect cornices based on edge detection in the rectified image. The third strategy is based on a combination of the former two strategies. We have implemented all three strategies using the following parameters: Threshold θ_g is used for inspecting the protrusion of a hypothesis, the two parameters δ and v are used for finding straight lines and arcs in an edge map, and the threshold θ_e is used for inspecting the position of a hypothesis in relation to the corresponding window.

4.2 *Cornice Detection Based only on 3D Geometry*

Firstly, we want to determine if the analysis of the 3D geometry of the point cloud corresponding to pixels in the search space is sufficient for cornice detection. For this purpose we discretised the search space into quadratic patches with the size $w_{ij} / 10 \times w_{ij} / 10$. For each patch we determine the average Euclidean distance of its corresponding 3D points from the façade plane. Significant ornaments in front of the wall will cover several patches, so that they will lead to a connected set of patches with an average distance above a threshold θ_g. We derive a binary image showing all patches with 3D points in front of the façade as foreground and use morphological operations to determine the patches forming the largest connected component. This component is regarded as hypothesis for a cornice and we test if its width w_c and height h_c in the coordinate reference system of the rectified image fulfils the two constraints

$$w_{ij} \leq w_c \leq 2 * w_{ij}, \quad h_c \leq ½\, h_{ij}. \tag{2}$$

If the test is successful, the foreground component is classified as a cornice. This method does not lead to a decision about the cornice type because the shape of a cornice can only be derived from its boundary, i.e. from detected edges.

4.3 *Cornice Detection Based only on Edges in the Image*

Here, we analyse extracted edges in the rectified image using (Canny 1986). For obtaining robust results, we suppress small details by low-pass filtering. In the resulting edge map, we search for straight edges and arcs (Fig. 4). We employ RANSAC to estimate the parameters of a straight edge or an arc from a set of edge pixels. All candidates for straight lines or arcs are accepted for further analysis if they are supported by a number v of connected pixels, i.e., these pixels are adjacent and have a distance to the straight edge or the arc of less than δ pixel.

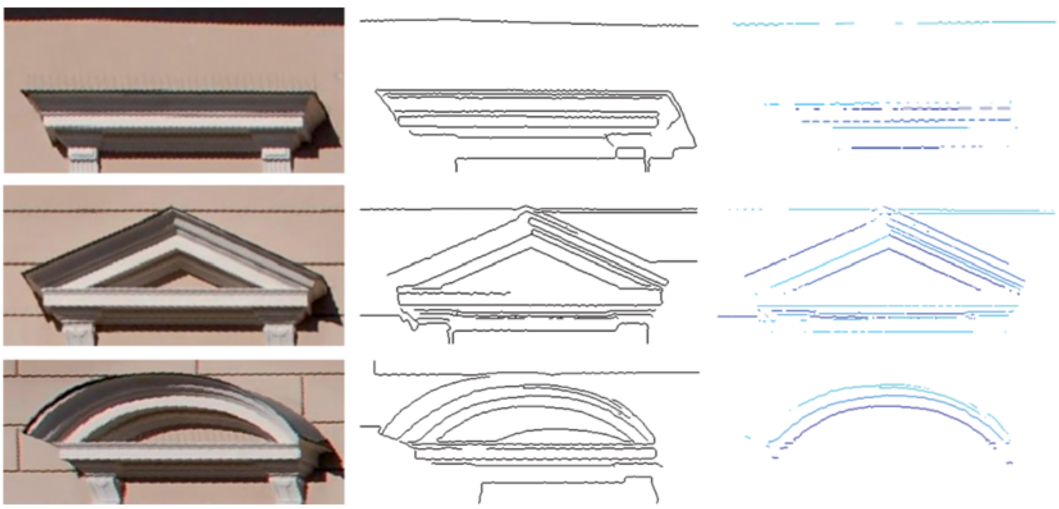

Fig. 4: RANSAC-based search for prominent straight edges and arcs. Left column: Image parts showing the three considered cornice types. Centre column: Edge maps obtained from Canny (1986). Right column: Detected straight edges (both upper rows) and arcs (bottom row).

We remove all straight lines with an absolute angle larger than 45° to the horizontal line. We choose such a large value, because most triangular cornices are obtuse-angled, i.e., the angles with the base must be less than 45°. Similarly, we remove all arcs without any supporting edge pixels in the neighbourhood of the arc's zenith.

We implemented a simple decision tree to decide if the extracted lines and arcs provide sufficient evidence for the presence of a cornice. We first check if we can find arcs whose centre lie at maximum θ_e pixels away from the vertical centreline of the detected window below the search space (Fig. 3). If there is also a horizontal line below an arc, this combination of the arc and the horizontal line is considered to be a hypothesis for an arch-shaped cornice. Otherwise, we check if there is a triangular cornice. This is true if we find two straight lines with equal absolute slope that are symmetrical to a vertical line and whose intersection is less than θ_e pixels away from the vertical centreline of the corresponding window. If there is also a horizontal line below this line pair, a hypothesis for a triangular cornice is generated. All possible hypotheses for triangular cornices are collected and analysed in an additional step. If no triangular cornice can be found, we finally check for a horizontal cornice. In this case, we must find at least two horizontal (parallel) edges.

If several line pairs or arcs fulfil the detection criterion of the decision tree, we first select all hypotheses which are symmetric with respect to the window's centreline, and we select the hypothesis whose top corner is highest above the window as the final result. Different cornice shapes do not have to be considered, because this hypothesis selection is performed in one of the steps of the decision tree. In all three cases we also determine the tolerance range depending on the perspective distortion of the original image, because we have to deal with small distortions concerning all objects in front of the façade plane (Fig. 3) since we only determined a homography instead of constructing an orthophoto.

4.4 *Combination of 3D-Geometry and Edges in the Image*

The performance of the cornice detection approach based on extracted edges depends on the values of the parameters θ_e, δ and v (rectilinearity and symmetry of the cornice to the window) while the performance of the approach based on 3D geometry depends on the allowed ranges of the cornice's width w_c and height h_c and on the value of the parameter θ_g (third dimension). First, we implemented the edge detection and carefully searched for proper values of the parameters θ_e, δ and v. However, we either obtained false detections caused by shadows and façade paintings, or missed the cornices.

Finally, we found that in a combination of both approaches, edge detection and 3D, we can set the range as proposed in (2) and tune the parameters θ_g, θ_e, δ and v almost without increasing the false detection rate (Fig. 5). We

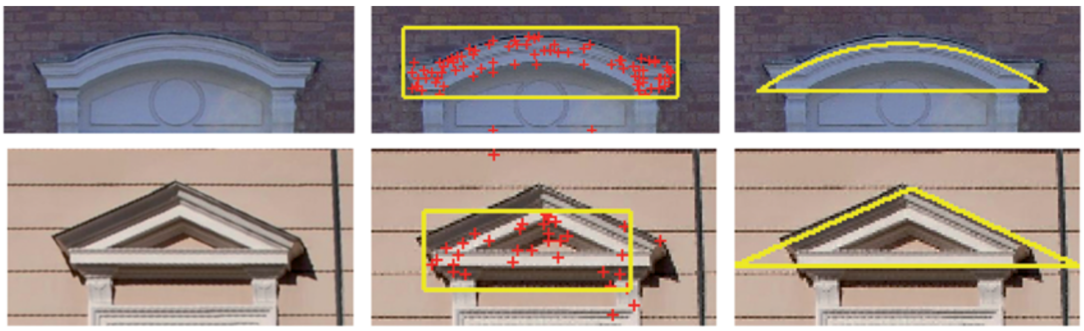

Fig. 5: Combining 3D geometry and edges in the image for cornice detection. Left column: image parts showing the two considered cornice types. Centre column: geometry-based cornice detection using a low threshold for acceptance. Right column: cornice hypothesis based on detected edges using low thresholds employing CANNY (1986). Both results are merged for reliable cornice detection.

included the test on the geometrical plausibility of a large structure located in front of the façade plane (section 4.2) in the edge-based decision tree for the cornice detection (section 4.3). In other words, each decision in the decision tree is confirmed by the additional check for a 3D structure in front of the façade. If the edge detection part finds a potential cornice, this object is then tested for its third dimension and accepted if it is higher than the threshold value θ_g above the wall.

5 Experiments

We tested all three variants of our approach with ten image sets, each showing one historic, medium-size façade from the city centres of the Bavarian cities Freising and Munich as well as the Franconian city Schwabach. Each image set consisted of at least seven images and the reconstruction of the point cloud was done using the approach of Mayer et al. (2012). Each point cloud contained about 9,000 3D points with a point density mostly between 20 and 50 points per m^2.

In our experiments we used the following parameter values. In the geometry-based approach, we have chosen the range for the cornice's width and height according to (2) and set $\theta_g = 0.05$ m. The latter value has been chosen to be slightly smaller than typical protrusions of cornices. For the edge-based approach, we smoothed the image with a small Gaussian filter with $\sigma = 1$ before calculating the gradients to handle the large radiometric variation in the façade images. The two thresholds $\theta_{c1} = 0.01$ and $\theta_{c2} = 0.1$ were used for non-maximum suppression and for tracing the edges with hysteresis thresholding, respectively, when employing (Canny 1986). The two parameters θ_e and δ were derived from the width w_{ij} of the window and under consideration of the image resolution. This yielded $\theta_e \approx 0.1$ m and $\delta = 1 – 2$ pixel. Additionally, the parameter v was chosen $v = w_{ij} / 2$. When combining both approaches, we changed the values of the parameters to $\theta_g = 0.03$ m, $\theta_e \approx 0.3$ m, $\delta = 3 – 5$ pixels and $v = w_{ij} / 3$.

The parameters for all three variants of our approach were chosen to avoid false detections. As a consequence, we did not recognize all existing cornices. The best results were obtained when combining information of the 3D geometry with the detected 2D edges (section 4.4). The detection results for triangular and arch-shaped cornices were significantly better than for horizontal cornices. We think that this is due to the characteristic shape of both cornice types while horizontal lines are often not reliable to distinguish horizontal cornices from horizontal shadows or textures.

In Tab. 1 we summarize the results of our experiments. The ten façades found in the different datasets contain a total of 95 windows in upper floors, i.e., 95 search spaces for cornice detection. In this experiment we did not consider shop windows or windows of cellars and dormers, because we had restricted the solution to the upper floors in one of the preprocessing steps and because usually these windows are not crowned by cornices. Six-

Tab. 1: Detection rates for the proposed of cornice detection regarding the three variants of our approach. In the first experiment (section 4.2), we only analysed the 3D geometry and did not explicitly classify the type of cornice. In the second experiment (section 4.3), we only analysed detected edges and automatically classified the type of cornice. In the third experiment (section 4.4), we combined both approaches including the cornice type selection. There were no false positives.

	windows without cornices	cornices all	cornices arch-shaped	cornices triangular	cornices horizontal
Ground Truth	16	79	39	21	19
Geometry (4.2)	16 (100 %)	52 (66 %)	29 (74 %)	16 (76 %)	7 (37 %)
Edges (4.3)	16 (100 %)	50 (63 %)	32 (82 %)	18 (86 %)	0 (0 %)
Combination (4.4)	16 (100 %)	61 (77 %)	32 (82 %)	18 (86 %)	11 (58 %)

teen of the 95 windows were not crowned by a cornice. We never detected a false positive. The detection rate of the 79 cornices was valid for all three variants of our approach. In the geometry-based approach (section 4.2) we did not classify the shape of the detected cornice. For the evaluation, we manually determined the shape-specific detection rates.

Mostly a low density of the point cloud in the area of a cornice leads to a non-detection of the cornice. Furthermore, many horizontal lines had a very weak contrast with the façade wall, i.e., the edge detection should be set more sensitive for detecting horizontal cornices. Since we did not know the type of a cornice in advance, the cornice detection should probably be performed multiple times with different parameters. Consequently, the different results would have to be merged and analysed, e.g. according to their overlap. This will be part of future work.

In our experiments, we were also challenged by partial occlusions. Cornices of some images were occluded by wires, which did not disturb the recognition. Larger occlusions, e.g. by trees, did not occur in our experiments. We suspect that our algorithm is not robust enough concerning such disturbances. Yet, since we acquire several images, we expect to see each cornice completely at least in one image. Thus, we could merge the results for cornice detection from different views.

We show two results for the combined approach with detected cornices in Fig. 6. Although the sizes and the positions of the detected windows were not always correct, the cornice detection often turned out to be satisfactory. It even returns good results for cornices above pairs of windows as shown at Fig. 6, bottom right.

6 Summary and Future Work

We have presented an approach for detecting cornices in rectified façade images and point clouds. The rectified façade images are derived from images together with highly precise medium-density point clouds. After detecting the façade plane and the windows we

Fig. 6: Results of cornice detection for façades from Freising (top row) and from Schwabach (bottom row) in Germany. All images were manually rectified. The image at top left is limited to the upper floors and shows the detected cornices in red and 3D points in front of the façade in blue. The other three images show the whole façade with detected windows in upper floors and cornices in red. Note that the cornices above the twin windows in the image at bottom right were successfully detected, because the window detection returned one window instead of two.

define search spaces for cornices crowning the windows. We consider three different cornice types: horizontal, triangular and arch-shaped. They are the most frequent types in European cities. In the experiments we obtained the best results if we considered information about 3D geometry together with straight edges or arcs detected in images.

Cornice detection could be improved by using denser point clouds. Since we employed the reconstruction approach of Mayer (2012), we only obtained a medium-density point cloud. If we used a dense matching approach such as semi-global matching (Hirschmüller 2008), we could generate much denser point clouds, having one 3D point for each pixel.

A further improvement would be the full automation of our approach, i.e., by an integration of further steps of the automatic façade extraction by Nguatem et al. (2012). This can only be achieved if images of complete street blocks are acquired. Particularly, Nguatem et al. (2012) can be used to detect balconies and oriels, which make the detection of the façade plane more difficult. We should also avoid a rectification of the image using a homography, because 3D structures in front of or behind the façade plane are projected to wrong positions in the image. If several façades of a street block have to be analysed, we also need to employ a façade segmentation, e.g., after Wendel et al. (2010). The use of orthophotos will eliminate this problem in the future. It is also planned to analyse several rectified views and merge their results. Additionally, further cornice styles like the shell-shaped cornices of Rococo will be considered for the future work.

Acknowledgements

The authors would like to thank their colleagues Jan Bartelsen and William Nguatem for support when reconstructing the scene by using the software of Mayer et al. (2012). Furthermore, we would like to thank the anonymous reviewers for their critical remarks and the editor Franz Rottensteiner for many suggestions which have helped to considerably improve this paper.

References

Akaike, H., 1973: Information Theory and an Extension of the Maximum Likelihood Principle. – **2nd** International Symposium on Information Theory: 267–281.

Becker, S. & Haala, N., 2008: Integrated LIDAR and Image Processing for the Modelling of Building Facades. – PFG – Photogrammetrie, Fernerkundung, Geoinformation **2008** (2): 65–81.

Canny, J., 1986: A Computational Approach to Edge Detection. – IEEE Transactions on Pattern Analysis and Machine Intelligence **8** (6): 679–698.

Fischler, M.A. & Bolles, R.C., 1981: Random Sample Consensus: A Paradigm for Model Fitting with Application to Image Analysis and Automated Cartography. – Communications of the ACM **24** (6): 381–395.

Gröger, G., Kolbe, T.H., Nagel, C. & Häfele, K.-H. (eds.), 2012: OGC City Geography Markup Language (CityGML) Encoding Standard. Version 2.0.0, OGC Doc. No. 12-019, Open Geospatial Consortium 2012.

Hirschmüller, H., 2008: Stereo Processing by Semi-global Matching and Mutual Information. – IEEE Transactions on Pattern Analysis and Machine Intelligence **30** (2): 328–341.

Hohmann, B., Krispel, U., Havemann, S. & Fellner, D., 2009: CityFit: High-Quality Urban Reconstruction by Fitting Shape Grammars to Images and Derived Textured Point Clouds. – The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences **38** (5/W1).

Jahangiri, M. & Petrou, M., 2008: Fully Bottom-Up Blob Extraction in Building Facades. – **9th** International Conference on Pattern Recognition and Image Analysis.

Koepf, H. & Binding, G., 2005: Bildwörterbuch der Architektur. – Fourth Edition, 678 p., Alfred Kröner Verlag, Stuttgart.

Leberl, F., Meixner, P., Wendel, A. & Irschara, A., 2012: Automated Photogrammetry for Three-Dimensional Models of Urban Spaces. – Optical Engineering **51** (2): 1–12.

Lee, S.C. & Nevatia, R., 2004: Extraction and Integration of Window in a 3D Building Model from Ground View Images. – IEEE Conference on Computer Vision and Pattern Recognition **II:** 113–120.

Mayer, H., 2007: 3D Reconstruction and Visualization of Urban Scenes from Uncalibrated Wide-Baseline Image Sequences. – PFG – Photogrammetrie, Fernerkundung, Geoinformation **2007** (3): 167–176.

Mayer, H., Bartelsen, J. Hirschmüller, H. & Kuhn, A., 2012: Dense 3D Reconstruction from Wide Baseline Image Sets. – Outdoor and Large-Scale Real-World Scene Analysis – 15th International Workshop on Theoretical Foundations of Computer Vision, Lecture Notes in Computer Science **7474:** 285–304.

Nguatem, W., Drauschke, M. & Mayer, H., 2012: Finding Cuboid-Based Building Models in Point Clouds. – The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences **39** (B3): 149–154.

Reznik, S. & Mayer, H., 2008: Implicit Shape Models, Self-Diagnosis, and Model Selection for 3D Facade Interpretation. – PFG – Photogrammetrie, Fernerkundung, Geoinformation **2008** (3): 187–196.

Ripperda, N., 2008: Grammar Based Facade Reconstruction Using rjMCMC. – PFG – Photogrammetrie, Fernerkundung, Geoinformation **2008** (2): 83–92.

Simon, L., Teboul, O., Koutsourakis, P. & Paragios, N., 2011: Random Exploration of the Procedural Space for Single-View 3D Modeling of Buildings. – International Journal of Computer Vision **93** (2): 253–271.

Šochman, J., 2007: Evaluation of the AdaBoost IPM. – Unpublished, internal note of project "eTraining for Interpreting Images of Man-Made Scenes (eTRIMS)".

Wendel, A., Donoser, M. & Bischof, H., 2010: Unsupervised Facade Segmentation Using Repetitive Patterns. – **32**nd DAGM conference on Pattern Recognition: 51–60.

Wenzel, S. & Förstner, W., 2012: Learning a Compositional Representation for Facade Object Categorization. – The International Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences **I** (3): 197–202.

Wonka, P., Wimmer, M., Sillion, F. & Ribarsky, W., 2003: Instant Architecture. – ACM Transactions on Graphics **22** (4): 669–677.

Zeng, G., Gan, R. & Zha, H., 2011: Building Facade Interpretation Exploiting Repetition and Mixed Templates. – First Asian Conference on Pattern Recognition: 652–656.

Addresses of the Authors:

Dipl.-Ing. Wolfgang Brandenburger, Amt für Geoinformationswesen der Bundeswehr, Dezernat III 1 (1) Geländeanalyse, Frauenberger Straße 250, D-53879 Euskirchen; Tel: +49-2251-953-4637, e-mail: wolfgangbrandenburger@bundeswehr.org

Dr.-Ing. Martin Drauschke, Univ.-Prof. Dr.-Ing. Helmut Mayer, Bundeswehr University Munich, Institute of Applied Computer Science, D-85577 Neubiberg, Germany; Tel: +49-89-6004-3454, -3429, Fax: -3906, e-mail: {martin.drauschke},{helmut.mayer}@unibw.de

Manuskript eingereicht: Mai 2013
Angenommen: Juli 2013