A Cascaded Classifier Approach to Window Detection in Facade Images

M. Neuhausen, A. Martin, M. Obel, P. Mark and M. König

Ruhr-University, Germany, Bochum
E-mail: marcel.neuhausen@ruhr-uni-bochum.de

Abstract -

A major part of recent developments in civil engineering in the urban context evolved around building and city models. Especially for a precise risk assessment of damages to existing buildings induced by ground movements, accurate models are inevitable. Beside the shape of a building, the focus is also on components compromising a building's stiffness. Particularly, by including wall openings such as windows into risk analyses, these can be improved to provide more reliable predictions.

However, most publicly available data sources only provide simple block models of existing buildings sometimes extended by roof shapes. As a consequence, any information concerning the windows of a building must be integrated into the model using other data sources. Whereas numerous approaches address the refinement of building shapes, their windows and other components are commonly disregarded. Although cascaded classifiers already turned out to yield good results in general and applying them to window detection seems promising, such approaches are yet insufficient to reliably extend building models. Drawing on previous findings, we present an approach to window detection in facade images satisfying the needs of risk assessment analyses. Our detection system combines a soft cascaded classifier consisting of thresholded Haar-like features with a sliding window detector extracting image patches for classification. The soft cascaded design improves the detection rate over previously made approaches while coincidentally reducing the amount of required features. Further, we evaluate the effect of a rectified dataset on the classification results compared to its counterpart with images taken from varying angles.

Keywords -

Window Detection; Cascaded Classifier; Urban Environments; Building Classification; Building Reconstruction

1 Introduction

Models of buildings are indispensable in nowadays civil engineering in the urban context. For several tasks in urban planning and inner city construction projects accurate models of existing buildings are a prerequisite. In particular, this expresses in risk assessment concerning damages to buildings induced by ground movements which may be caused by earthquakes or tunneling projects. Reliable analyses predicting the risk of damages demand detailed information about a structure's stiffness. Beside the pure shape of a building, risk analyses also consider certain facade components as they possess more profound information about a structure's condition. Especially the ratio of wall openings to facade is taken into account since openings highly impair the stiffness of a structure [9]. However, wall openings are commonly not included in publicly available 3D building models. Those requested from land registry offices usually comprise coarse block models of buildings. Some models may be extended by simplified roof shapes but in general they lack any further detail as shown in figure 1. Except for a few manually reconstructed buildings, those provided by online map services like OpenStreetMaps and Google Earth neither offer more detailed information. For that reason, such models have to be enriched with additional information about wall openings beforehand to satisfy the requirements of risk analyses.

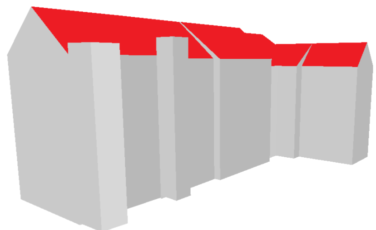

Figure 1. 3D block model of a group of existing buildings provided by the land registry office of Düsseldorf.

For deriving wall openings from a facade, its windows qualify best as they commonly account for almost all openings. Nevertheless, a personal inspection of each building within an urban area or along a tunneling alignment to determine the window-to-facade ratio is time consuming and expensive. Several approaches towards a computer-aided detection of windows in facades have already been made. Many of them rely on assumptions and restrictions to the buildings' architecture which make them impractical for general usage. To detect windows in facades automatically without any prior knowledge about the investigated build-

(a) (b)

Figure 2. (a) Detected window in non-rectified image. Area of the detection rectangle (red) and the window are not coincident. (b) Detected window in rectified image. Nearly coincident areas of detection rectangle (red) and window.

ing, Ali et al. [4] apply the Viola-Jones framework [15] to images of building scenes. Despite a stated detection rate which is insufficient for our purpose, it indicates that cascaded classifiers yield promising results for window detection tasks. Based on their findings, this paper focuses on improving the detection rate and the accuracy regarding the windows area. We substitute the classifier as proposed by Viola and Jones [15] with a soft cascaded classifier [7]. This provides better generalization for high variation of the training set which results in an improved detection rate and higher precision while coincidentally employing less features. Apart from that, we keep the original detection algorithm of the Viola-Jones framework which scans the entire image since the soft cascaded classifier is even slightly faster so that there is no need for improvements. Furthermore, we rectify the facade images before detection to obtain a more precise congruence of the window areas with the rectangular detections of our approach (see fig. 2). Restricting to rectangular windows also eases the training and classification process and consequently increases the detection rate.

These improvements increase the quality of the detections to a satisfactory level in terms of the aforementioned risk assessment. The results of our window detection system as proposed in this paper can, thus, be incorporated into the building model to facilitate subsequent analyses.

2 Related Work

In recent years, window detection in facades has become prominent in a diversity of research fields. The area of application ranges from tourist guidance on mobile devices where buildings are recognized by characteristic window patterns over 3D city modeling to deformation analysis [4]. However, in literature window detection is commonly referred to as a part of building reconstruction.

A wide range of approaches concerning several facets of building reconstruction have already been made. Most of them focus on improving the model's shape disregarding its components. A survey concerning such approaches by Brenner [1] focuses on aerial imagery and laser scans as input data. This facilitates data acquisition with appropriate effort even for entire cities. Information about large structures like buildings and streets is well extractable which enables the reconstruction of cities as simple block models. An advanced reconstruction of a building, especially by its facade components, is only feasible to a very limited extent due to the bird's-eye perspective.

Haala and Kada [10] additionally include approaches based on terrestrial laser scans into their survey which provide accurate point clouds of existing buildings from street level. Beside generating geometry from these measurements [18], this kind of data enables the detection of certain facade components such as windows by identifying no-measurement areas in the facade plane [17]. Methods relying on terrestrial laser scans offer simple ways to detect windows in facades, but by now such data is not available regionwide. Albeit the StreetMapper proposed by Barber et al. [3] eases the acquisition, it is too costly and involves unreasonable effort.

On the contrary, ground view images of facades are publicly available from web services like Google Street View or can easily be gathered even for larger areas. In preliminary work Neuhausen et al. [8] discuss window detection approaches using street level images. They distinguish contemporary approaches by their shared strategies into three categories: Grammar-based, Image processing, and machine learning. Grammar-based approaches highly rely on several assumptions to a facade's appearance and prior knowledge about the architecture and its composition. Ripperda [11] and Ripperda and Brenner [12] divide facades into their components by applying a simple grammar leveraging symmetries and repetitions in facade images. In contrast, Teboul et al. [13] based their approach on a detailed shape grammar and combined it with a model of semantic relationships between certain facade elements. The definition of a proper set of grammar rules and the modeling of relationships is non-trivial and requires some expert knowledge about the expected architecture. Such approaches have to be manually adjusted whenever the architectural conditions change. This results in a limitation of a general application. As grammars allow several decompositions of a building's facade the most probable subdivision has to be found by means of sampling methods or parsing algorithms. This dramatically raises the methods' complexity with the number of rules in the grammar. By this reason, rule sets have to be kept small which complicates a detailed modeling. Similar to grammar-based approaches, image processing methods also make assumptions about the alignment of windows on the facade. Lee and Nevatia [2] presume windows to be aligned in a grid-like manner on the facade as usually seen at big office buildings. For window detection they project horizontal and vertical edges of rectified facade images into histograms. These are then superimposed to find window hypotheses, refined in later steps, at the peaks' locations. Pursuing this approach, Meixner et al. [14] found that it provides good results on highly regular facades but fails as soon as facades become more complex. Asymmetric patterns of windows as well as facade extensions like balconies corrupt the reliability of this method. As the facades' appearance not only alters between countries but also between city districts and, moreover, most facades in the urban area can be expected to be complex, these approaches will often fail and are, thus, not universally applicable. To avoid the aforementioned limitations, windows have to be detected by their inherent characteristics instead of their relations among each other or further facade elements. Machine learning algorithms emerge beneficial as they can be trained to classify windows by image features which are directly linked to these characteristics. Haugegard et al. [6] emphasize this by deploying a support vector machine (SVM) as classifier to identify windows by their edges. The set of features used for classification, though, has to be chosen manually. A boosted cascaded classifier as proposed in the Viola-Jones framework [15] overcomes this by autonomously drawing those features from a predefined pool which qualify best for a certain task. Furthermore, these usually outperform most monolithic classifiers such as SVMs [16]. In their approach, Ali et al. [4] apply this framework to window detection and investigate the quality of its detections. The resulting detection rates are insufficient to reliably estimate the window-to-facade ratio of a building but, nevertheless, cascaded classifiers emerge to be promising to further research.

In the following we give an insight into the concept of our detection system (see sec. 3.1) which is based on these findings. We outline the classifier used and the detection algorithm in section 3.2 and section 3.3. In section 4.1 we compare our system to the approach of Ali et al. [4]. In further experiments we focus on improving the detection rate. For this purpose, we contrast the classification system of the previous experiment with one operating on rectified facade images (see sec. 4.2). Finally, we conclude our findings and provide an outlook on further work to be done in section 5.

3 Proposed Window Detection System

Window detection in facade images is still demanding and not sufficiently solved yet. The detection proves to be challenging for multiple reasons. Most of a window is made of a transparent pane which restricts detection to window frame. Additionally, windows occur in different sizes and shapes which further increases the difficulty of the detection task due to a high variation in the frames' appearance. Both complicate the manual choice of an adequate set of image features as it is uncertain which subset characterizes the entire class of windows best. While training, boosted cascaded classifiers choose those features from a predefined pool which perform best on a certain dataset. They overcome the issue of a manual selection if the training dataset is well constructed so that it depicts high intraclass variation. Moreover, they generally yield better results than monolithic classifiers. For these reasons, boosted cascaded classifiers seem to be promising for our window detection system.

3.1 Concept

Since boosted cascaded classifiers proofed to be advantageous compared to other classification methods, we build our detection system as outlined in figure 3 on the approach of Ali et. al [4] which uses the Viola-Jones framework. Considering their findings, we constrain the input data to be rectified semi-automatically. That eliminates an unnecessary increase in the windows' variation since equal windows no longer appear to be different due to distortion. Nevertheless, the intraclass variety remains quite high. We substitute the classifier used in the Viola-Jones framework by the soft cascaded classifier proposed by Bourdev and Brandt [7] as it is more robust regarding high variability in the set of positive training samples. Furthermore, this method commonly offers a better detection rate while relying on less features. In section 3.2 we describe its functionality in more detail. The soft cascaded classifier we use in our approach is even slightly faster than the one applied in the Viola-Jones-framework due to the lower number of features that have to be evaluated. An optimization of the enclosing detection algorithm is therefore not necessary. We keep the naive sliding window detector (see sec. 3.3) of the Viola-Jones framework which scans an image via small subwindows sliding across the entire image and passes the subjacent image patches to the classifier. The classifier which is trained on a set of rectified window samples and random non-windows then obtains a binary classification for each patch and returns them to the detector. Positive responses are finally merged where appropriate and returned as output of our system. Figure 3 illustrates the described interactions of the detectors' components.

3.2 Soft Cascaded Classifier

As an improvement over the Viola-Jones framework, Bourdev and Brandt [7] proposed the soft cascaded classifier approach which also relies on an set of weighted weak

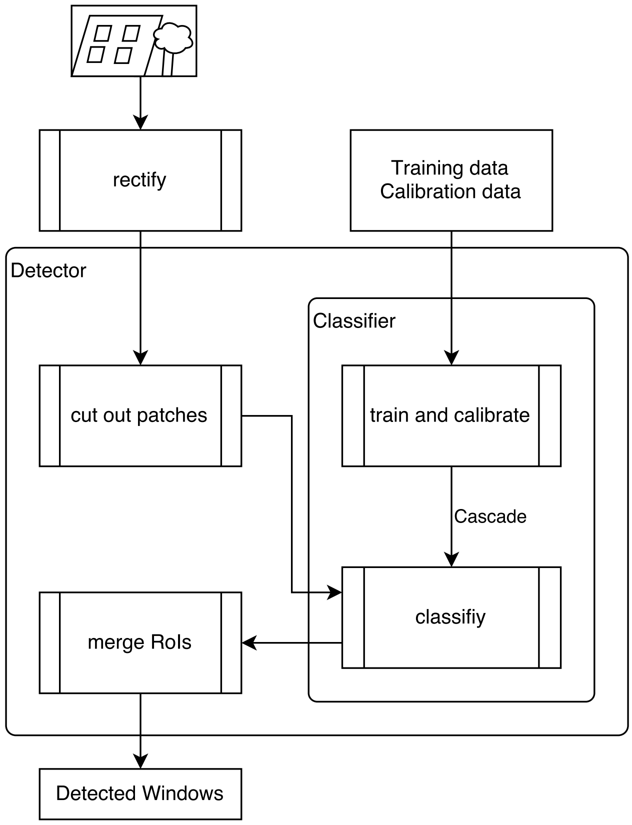

Figure 3. Concept of the proposed detection system. Images are rectified and passed to the detector. The detector cuts out image patches which are classified by a trained and calibrated cascade. Overlapping regions of interest (ROIs) of positive classifications are merged. Resulting ROIs are returned as detected windows.

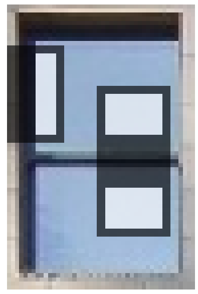

Figure 4. Two Haar-like features at different positions in the image patch, each corresponding to a weak classifier.

classifiers $\alpha h(x)$, where α is the weight and $h(x)$ denotes the classification function of the weak classifier for a given sample x. Each $h(x)$ classifies barely better than randomly guessing. By connecting them in series a strong classifier with high detection rate emerges. Equal to the Viola-Jones framework the weak classifiers utilized here are Haar-like features whose responses are thresholded to enable binary classification. The thresholds are set independently for each feature by the algorithm given by Viola and Jones [15]. According to this, each weak classifier consists of a thresholded Haar-like feature of a certain size at a certain relative position in the samples' image patches (see fig. 4).

For training a dataset $\{(x_1, y_1, w_1)...(x_N, y_N, w_N)\}$ has to be constructed with positive and negative samples x, their corresponding true labels y and a weight w indicating the importance to classify the sample correctly for increasing the classifier's quality. In the beginning all weights are set to be equal. While training, those weak classifiers performing best on the training dataset with respect to the samples' weights w are iteratively drawn from a pool and added to the strong classifier's set. The weight of the weak classifier within the set is determined according to its training error. The lower the error is, the higher is the classifier's weight. As covering the infinite class of non-windows by a finite set of negative samples is impossible, it can only be approximated. For this reason, after adding a weak classifier to the set new, negative samples are bootstrapped which are yet misclassified by the current strong classifier and added to the existing dataset. The weight w_i of each sample of the set is then adapted corresponding to its classification result so that currently misclassified samples increase in importance and vice versa. Since this promotes the selection of weak classifiers obtaining the correct class of currently misclassified samples, the focus of the emerging strong classifier is shifted towards these samples. This further improves the strong classifier to correctly classify a wide spectrum of samples.

Out of this a strong classifier emerges that can already be used for classification. Therefore a weighted majority vote of all holded weak classifiers is done:

$$\sum_t^T \alpha_t h(x)_t \geq \frac{1}{2} \sum_t^T \alpha_t$$

However, evaluating all weak classifiers on a sample is very time consuming and unnecessary unless the sample is positive. Arranging the weak classifiers within the strong classifier in a cascaded layout highly decreases the required evaluations. A subsequent weak classifier is then only evaluated if its predecessors classified the sample positive. Otherwise the evaluation is terminated prematurely and the sample is classified negatively. This behavior is implemented by a sample trace which contains the partial sum of all already evaluated weak classifiers. After each evaluation, the sample trace is compared to a rejection threshold to determine whether the sample is further promoted through the cascade or rejected as negative sample. An illustration of this procedure is given in figure 5. In contrast to the Viola-Jones framework, each weak classifier represents a single stage within the cascade. The decision if a sample is rejected becomes softer as it does not only depend on the previous stage but on all stages previously evaluated. The order of the weak classifiers as well as the rejection thresholds highly affect the quality of the detection. A proper setting of those is done in a calibration step on an already trained classifier. Given a calibration dataset which is constructed like the training dataset but contains a new set of samples, the weak classifiers are reordered with regard to the performance on this set. For each sample in the set, its sample trace is maintained. Based on these traces the rejection threshold of each stage can be set so that each stage rejects as much negative samples as possible but only a certain percentage of positive samples. Once the classifier is calibrated, the stages are evaluated one after the other and each resulting partial sum is compared to the corresponding rejection threshold of the stage. When the sum drops below the threshold, the sample is discarded and marked as negative. If, elsewise, the sample passes through all stages, it is classified to be positive.

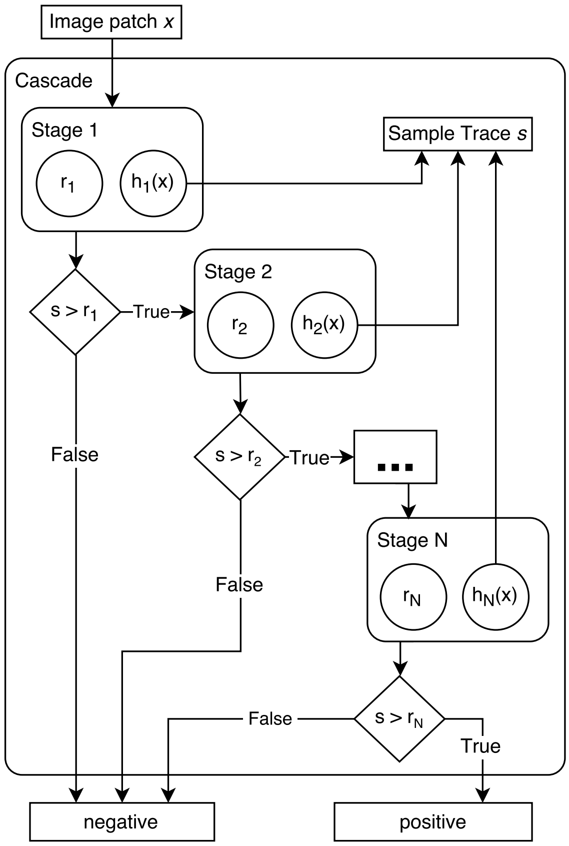

Figure 5. Schematic illustration of the cascaded classification procedure. The weak classifiers of each stage are only evaluated if the sample trace s is higher than the rejection threshold r_i of the preceding stage. A sample is positively classified if it passes through the whole cascade. Otherwise it is prematurely declined and negatively classified.

3.3 Sliding Window Detector

The classifier of our system accepts image patches and returns whether they mainly consist of a window or not. For finding windows in facade images, hence, a detector is required which passes relevant patches to the classifier and processes the returned decisions. Due to the cascaded structure of the classifier, negative patches are declined in early stages so that they can be processed very fast. For this reason there is no need for an optimized detection algorithm.

We use the naive sliding window detector proposed by Viola and Jones [15]. For a given image a rectangular subwindow slides across the entire image in multiple scales. The shifting step size depends on the particular scale of the subwindow. At each position the enclosed image patch is passed to a classifier. If the classifier resolves a patch to contain an object of interest, i.e. a window, the subwindow is memorized as a region of interest (ROI). Since the cascaded classifier is to some extend insensitive to changes in translation and scale of the object within the image patch, multiple ROIs may occur within a region around an object [15]. The memorized ROIs are reprocessed such that overlapping ones are averaged, resulting in a single ROI.

We specify the images passed to the detector to be rectified and completely filled by facade, so we can approximate expected window dimensions by the image's dimensions. Based on this, we define the minimal sliding subwindow dimensions to be half that size which corresponds to a starting scale of $s = 1.0$. After each run across the facade image the subwindow is scaled by a factor of 1.25. The step size of shifting starts with $\Delta = 1.0$ and is scaled after each run by $[s\Delta]$, where [] denotes a rounding operation. At each position the underlying image patch is passed to the classifier. To separately detect windows which are positioned close together we only merge positively classified ROIs which overlap at least by a factor of 0.75.

4 Experiments

In the following we present two experiments. In the first experiment, described in section 4.1, we run our detection system on non-rectified facade images. We compare our detection results to those of the approach of Ali et al. [4]. By contrasting the results we examine the improvements of the soft cascaded classifier over the Viola-Jones framework on the window detection task. In a second experiment (see section 4.2) we run our system on a set of rectified facade images. We then compare the results to our system of the non-rectified experimental setup to highlight the effect of rectified windows to the classification results.

4.1 Detection on Non-Rectified Images

To train our classifier analogous to Ali et al. [4] we manually cut out 1500 windows from the TSG-60 [20] and ZuBuD [5] image datasets. Both sets contain images of facades taken from different viewpoints and in varying illumination conditions. In contrast to Ali et al. [4] we discard the TSG-20 [19] dataset from training as it is a subset of TSG-60. Instead we take 900 positive samples from the TSG-60 dataset which were taken from a viewpoint of approximately 30° to the facades plane. Another 600 windows from various viewpoints are taken from the ZuBuD dataset. We also add the vertically flipped version of each window to our training dataset so that the set additionally covers windows taken from opposing viewpoints. Overall our training set contains 3000 positive samples and, initially, 3000 randomly chosen non-windows as negative samples. After training we calibrate our classifier on further 3000 samples of these datasets.

For evaluation we use the setup of Ali et al. [4] to compare their detection results to ours. Therefore we evaluate our detection system with the calibrated classifier on the same three distinct datasets. We use each 40 images of both, the TSG-20 and TSG-60 dataset. We also use 115 images of the ZuBuD Query Images which contain a subset of the facades shown in the ZuBuD dataset. All of the images considered for evaluation are taken from different viewpoints and illumination conditions than those of the training set.

As the exact position and size of the detected windows are relevant for subsequent risk analyses we apply the single window evaluation approach of Ali et al. [4]. According to this, a detection is marked as true positive only if the hypothesis is inside the manually labeled ground truth rectangle or has a maximal overlap of 5 pixels in each direction. Additionally, the hypothesis has to cover at least 75% of the underlying bounding box of the ground truth label. A hypothesis is marked as false positive if it covers less than 5% of a ground truth label. The detection results are summarized in table 1.

Dataset	True positive in %	False positive in %
TSG-20	69.1	7.7
TSG-60	60.3	8.9
ZuBuD	39.6	3.1

Table 1. Detection results of our system on non-rectified datasets.

For applying the single window evaluation with a coverage of at least 75% Ali et al. [4] stated true positive rates of 57%, 52%, and 30% on the TSG-20, TSG-60, and ZuBuD Query Images datasets, respectively. False positives were below 9% for all datasets. With similar false positive rates and true positive rates of 69.1% on TSG-20, 60.3% on TSG-60, and 39.6 on ZuBuD Query Images our detection system performs better on non-rectified facade images. By the substitution of the classifier we achieve an increase in the true positive rate up to 12% on the TSG datasets while keeping the false detections on the same level. Although ZuBuD constitutes a very challenging dataset because of many different viewpoints and uncommon buildings with unique window shapes such as churches, our system increases the true positive rate by 9% associated with a marginal increase in false positive rate by 1%. Nevertheless, in general these rates are far from being sufficient for precise risk assessment analyses as visible in the sample images in figure 6. Since we only exchanged the classification method, the results of this experiment demonstrate the improvement of the soft cascade compared to the Viola-Jones classifier. Thus, the soft cascade approach offers a better basis to further improve the detection rate.

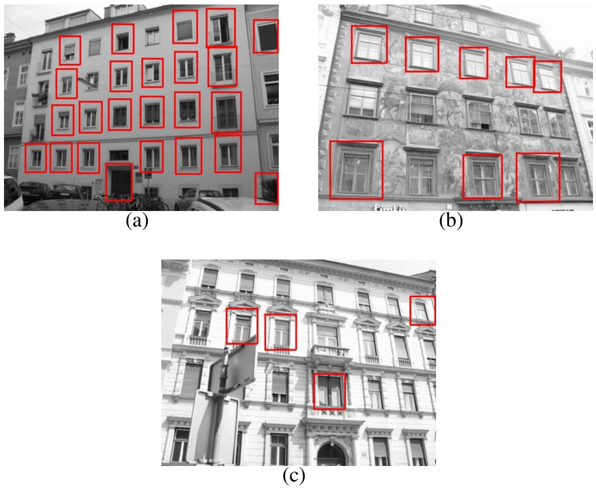

Figure 6. Detections on different facade types an viewpoints. (a) Good detection rate on simple facades with common windows. (b) Decreased performance on complex facade appearance. (c) Despite simple facade few detections due to viewpoint.

4.2 Detection on Rectified Images

In this experiment we compare the soft cascaded classifier's performance on non-rectified facade images to the performance on rectified images. For training the classifier which runs on rectified samples, we use the Ecole Centrale Paris Facades Database [13]. The dataset consists of 478 rectified facade images from various cities in Europe and the United States. We take 3000 randomly chosen windows from this and initially add the same amount of non-windows to the training set. For comparison we use the classifier already trained in the experiment in section 4.1 to run on non-rectified samples.

We evaluate the classifier for rectified samples on 3080 positive and 10000 negative image patches of the Ecole Centrale Paris Facades Database which were not shown while training. For the evaluation of the other classifier, we use 3102 positive and 10000 negative image patches equally distributed of the TSG-20, TSG-60, and ZuBuD Query Image datasets. We compare both classifiers by means of their receiver operating characteristic (ROC) curve shown in figure 7. Their curves are generated by plotting their detection rate against false positive rates while increasing the classification threshold from 0.5 up to 1.

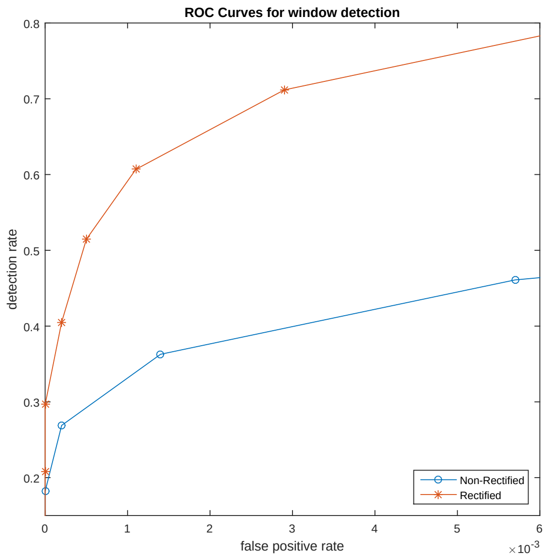

Figure 7. Comparative detection performances of the non-rectified and the rectified classifier by means of their ROC curves.

As the ROC curves of the classifiers clearly illustrate, the classifier for rectified images outperforms the one running on non-rectified images. Especially for false positive rates lower than 0.004% it provides reasonable detection rates of more than 71% whereas the non-rectified classifier only achieves around 40%. To obtain sufficient detection results very small false positive rates are a prerequisite since the detector scans hundred thousands of subwindows per image. Although the classifier for rectified facade images we trained in this experiment provides much better detection rates than the one for non-rectified images, its false positive rates are still too high for a precise prediction. By applying our detection system to rectified facade images at least a sufficient approximation to the real window-to-facade ratio is practicable as can be seen in the samples in figure 8.

(a) (b)

Figure 8. Detections on rectified facades. (a) Detections of windows close together may be shifted due to merging. Some windows are missed. (b) Even slightly occluded windows are detected. No detection of store windows.

5 Conclusion

For risk assessment of ground movement induced damages to existing buildings, analyses are made which require precise information about a structure's stiffness. Since publicly available data sources only provide block models of buildings, these have to be enhanced to facilitate analyses. As windows highly decrease the stiffness of a structure, we developed an approach to window detection in facade images based on the findings of Ali et al. [4]. Our approach consists of a sliding window detector in combination with a soft cascaded classifier.

In the experiments, we showed that this setup improves over the Viola-Jones framework. Additionally, we tested our classifier on rectified facade images which further improves the detection rate. The results our system achieved in the experiments are sufficient for analyses estimating a damage class for an existing building. Thereby many buildings along a potential tunneling alignment or within an earthquake area may automatically be discarded from further personal investigation as the risk of damages is very low. This highly reduces human effort for inspection of buildings resulting in time and cost savings.

However, the detection rate is still too low for a precise analysis. Some windows are slightly below the rejection threshold due to occlusions or bad illumination conditions and are thus declined by the classifier. To increase the detection rate future work may cover a post processing in which possible window positions are derived from the windows already detected. At these positions classifiers with slightly lower thresholds can be used to detect previously missed windows.

Acknowledgement

Financial support was provided by the German Research Foundation (DFG) in the framework of the Collaborative Research Center SFB 837 "Interaction Modelling in Mechanized Tunneling".

References

[1] Brenner C. Building reconstruction from images and laser scanning. International Journal of Applied Earth Observation and Geoinformation, 6(3–4):187–198, 2005.

[2] Lee S. C. and Nevatia R. Extraction and integration of window in a 3d building model from ground view images. In Proceedings of the Conference on Computer Vision and Pattern Recognition, pages 106–113. IEEE Computer Society, 2004.

[3] Barber D., Mills J., and Smith-Voysey S. Geometric validation of a ground-based mobile laser scanning system. ISPRS Journal of Photogrammetry and Remote Sensing, 61(1):128–141, 2008.

[4] Ali H., Seifert C., Jindal N., Paletta L., and Paar G. Window detection in facades. In Proceedings of the 14th Conference on Image Analysis and Processing, pages 837–842. IEEE Computer Society, 2007.

[5] Shao H., Svoboda T., and Van Gool L. Zubud – zürich buildings database for image based recognition. Technical Report 260, Computer Vision Laboratory, Swiss Federal Institute of Technology, 2003.

[6] Haugeard J., Philipp-Foliguet S., Precioso F., and Lebrun J. Extraction of windows in facade using kernel on graph of contours. In Proceedings of the 16th Scandinavian Conference on Image Analysis, pages 646–656, 2009.

[7] Bourdev L. and Brandt J. Robust object detection via soft cascade. In Proceedings of the Conference on Computer Vision and Pattern Recognition, pages 236–243. IEEE Computer Society, 2005.

[8] Neuhausen M., Koch C., and König M. Image-based window detection - an overview. In Proceedings of Workshop of the European Group for Intelligent Computing in Engineering, 2016.

[9] Obel M., Neuhausen M., Schindler S., König M., and Mark M. Settlement risk assessment by means of categorized surface infrastructure. In Proceedings of the International Conference on Structural Analysis of Historical Constructions, pages 511–518, 2016.

[10] Haala N. and Kada M. An update on automatic 3d building reconstruction. ISPRS Journal of Photogrammetry and Remote Sensing, 65:570–580, 2010.

[11] Ripperda N. Grammar based facade reconstruction using rjmcmc. Photogrammetrie Fernerkundung Geoinformation, 2:83–92, 2008.

[12] Ripperda N. and Brenner C. Application of a formal grammar to facade reconstruction in semiautomatic and automatic environments. In Proceedings of the AGILE Conference on Geographic Information Science, 2009.

[13] Teboul O., Simon L., Koutsourakis P., and Paragios N. Segmentation of building facades using procedural shape priors. In Proceedings of the Conference on Computer Vision and Pattern Recognition, pages 3105–3112. IEEE Computer Society, 2010.

[14] Meixner P., Leberl F., and Brédif M. Interpretation of 2d and 3d building details on facades and roofs. In Proceedings of the Conference on Photogrammetric Image Analysis, pages 137–142. ISPRS, 2011.

[15] Viola P. and Jones M. J. Robust real-time face detection. International Journal of Computer Vision, 57(2):137–154, May 2004.

[16] Lienhart R., Kuranov A., and Pisarevsky V. Empirical analysis of detection cascades of boosted classifiers for rapid object detection. In B. Michaelis and G. Krell, editors, Pattern Recognition. DAGM 2003. Lecture Notes in Computer Science, volume 2781, pages 297–304. Springer, Berlin, Heidelberg, 2003.

[17] Becker S. and Haala N. Refinement of building facades by integrated processing of lidar and image data. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 36:7–12, 2007.

[18] Pu S. and Vosselman G. Knowledge based reconstruction of building models from terrestrial laser scanning data. ISPRS Journal of Photogrammetry and Remote Sensing, 64(6):575–584, 2009.

[19] TSG-20: Tourist sights graz image database. http://dib.joanneum.at/cape/TSG-20/.

[20] TSG-60: Tourist sights graz image database. http://dib.joanneum.at/cape/TSG-60/.