BUILDING FAÇADE SEPARATION IN VERTICAL AERIAL IMAGES

P. Meixner, A. Wendel, H. Bischof and F. Leberl

Graz University of Technology, Institute for Computer Graphics and Vision, Inffeldgasse 16/II, Graz, Austria
{meixner, wendel, bischof, leberl}@ icg.tugraz.at

ISPRS Commission III, WG III/4 - Complex Scene Analysis and 3D Reconstruction

KEY WORDS: Vertical Aerial Images, Façade Separation, Semantic Building Interpretation, Building Contour

ABSTRACT:

Three-dimensional models of urban environments have great appeal and offer promises of interesting applications. While initially it was of interest to just have such 3D data, it increasingly becomes evident that one really would like to have interpreted urban objects.

To be able to interpret buildings we have to split a visible whole building block into its different single buildings. Usually this is done using cadastral information to divide the single land parcels. The problem in this case is that sometimes the building boundaries derived from the cadastre are insufficiently accurate due to several reasons like old databases with lower accuracies or inaccuracies due to transformation between two coordinate systems. For this reason it can happen that a cadastral boundary coming from an old map is displaced by up to several meters and therefore divides two buildings incorrectly.

To overcome such problems we incorporate the information from vertical aerial images. We introduce a façade separation method that is able to find individual building façades using multi view stereo. The purpose is to identify the individual façades and separate them from one another before on proceeds with the analysis of a façade's details. The source was a set of overlapping, thus "redundant" vertical aerial images taken by an UltraCam digital aerial camera.

Therefore in a first step we determine the building block outlines using the building classification and use the height values from the Digital Surface Model (DSM) to determine approximate "façade quadrilaterals". We also incorporate height discontinuities using the height profiles along the building outlines to enhance our façade separation. In a next step we detect repeated pattern in these "façade images" and use them to separate the façades respectively building blocks from one another.

We show that this method can be successfully used to separate building façades using vertical aerial images with a very high detection rate of 88%.

1. MOTIVATION

Accurate and realistic three-dimensional models of urban environments are increasingly important for applications like virtual tourism, city planning, internet search and many emerging opportunities in the context of ambient intelligence. Applications like Bing Maps or Google Earth are offering virtual models of many major urban areas worldwide. Initially such data were just used for visualization purposes, but this is on the way to change. On the horizon are urban models that consist of semantically interpreted objects. In its most sophisticated form, each building, tree, street detail, bridge and water body is modelled in three dimensions, details such as windows, doors, façade elements, sidewalks, manholes, parking meters, suspended wires, street signs etc. exist as separate objects.

To be able to interpret buildings we have to split a visible whole building block into its different single buildings. Usually this is done using cadastral information to divide the single land parcels. The problem in this case is that sometimes the building boundaries derived from the cadastre may be insufficiently accurate, for example due to old databases with lower accuracies or inaccuracies due to transformation between two coordinate systems. For this reason it can happen that a cadastral boundary coming from an old map is displaced by up to several meters (Feucht 2007) and therefore divides two buildings incorrectly (see Figure 1).

Figure 1. Overlaying the cadastral map (depicted in red) over the true orthophoto shows displacements of the cadastral boundaries versus the photography.

Therefore we incorporate the information from vertical aerial images. We employ a method introduced by Wendel et al. (2010) that is able to separate building façades in single images. Separate façades can then get analysed for their details. In this previous project, the source material consisted of a set of overlapping thus redundant images using a moving vehicle and calibrated automated cameras.

In the current project we adapt their method to vertical aerial images in the hope to increase the accuracy of a building block separation beyond that obtainable from previous approaches. We determine the building block outlines using the building classification and use the height values from the Digital Surface Model (DSM) to determine the approximated "façade quadrilaterals". We also incorporate height discontinuities using the height profiles along the building outlines to enhance our façade separation. In a next step we detect repeated pattern in these "façade images" and use them to separate the façades respectively building blocks from one another. As the major contribution of this paper, we show that it is possible to separate façades by just using vertical aerial images and height information derived from those images. We also show that the achieved accuracies are close to those available from street side images despite of a far lower geometric resolution of the aerial data compared to street side images.

We have evaluated the method for a test area that covers 400m x 400m near the core of the city of Graz (Austria) with 186 different buildings consisting of 65 major building blocks. We show that the proposed method can be successfully used to separate building façades using vertical aerial images with a detection rate of 88%. Figure 2 illustrates one result of the proposed façade separation approach.

Figure 2. Result of façade separation for one building block from the test area. As one can notice all buildings were separated correctly. Separations marked in red correspond to splits based on building height, while those marked in green are the result of repetitive pattern analysis.

2. SOURCE DATA

In order to produce good results one needs (a) a Digital Surface Model (DSM) with well-defined building roof lines to avoid ragged building edges, as well as (b) a precise classification image from the test area. We want to present an overview of these two products that are derived fully automatically from vertical aerial images. Figure 3 shows an example of these two products covering the test area.

The Digital Surface Model is created by "dense matching". The input consists of the triangulated aerial photographs. In the process, one develops point clouds from subsets of the overlapping images and then merges (fuses) the separately developed point clouds of a given area. The process is described by Klaus (2007). The postings of the DSM and DTM are at two pixel intervals, thus far denser than traditional photogrammetry rules would support. The conversion of the surface model DSM into a Bald Earth Digital Terrain Model DTM is a post-process of the dense matching and has been described by Zebedin et al. (2006).

Any urban area of interest is being covered by multiple color aerial images. These can be subjected to an automated classification to develop information layers about the area. We consider these to be an input into our characterization procedures. The classification approach used here has been described by Zebedin et al. (2006). However, classification and segmentation methods are topics of intense research. For example, Kluckner et al. (2009) have proposed Random Forests as an alternative novel method with good results specifically interpreting urban scenes imaged by the UltraCam digital aerial camera.

Standard classification of 4-channel digital aerial photography typically leads to 7 separate areas for buildings; grass; trees; sealed surfaces; bare Earth; water; and other objects shown as "unclassified". The unclassified areas may show lamp posts, cars, buses, people etc.

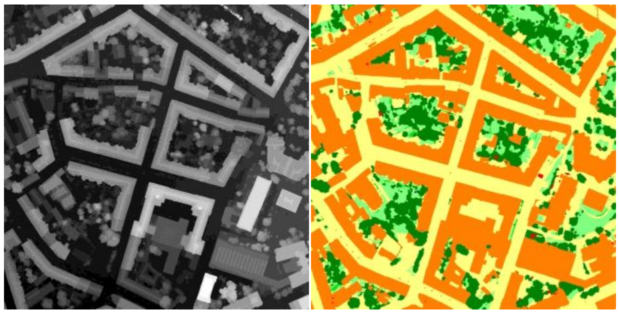


Figure 3. (a) Digital surface model from the test area (b) Classification image (orange: buildings, yellow: sealed surfaces, turquoise: bare earth, light green: grass, dark green: trees, blue: water, unclassified: red)

3. PROPOSED METHOD

This section describes how we separate buildings from vertical aerial images by applying a façade separation method introduced by Wendel et al. (2009). The proposed method consists of three steps. First we extract the building block outlines using the classification image. To generate the necessary straight lines we apply a recursive line simplification scheme on the building block footprints. In a second step the height values coming from the DSM are assigned to the extracted building block outlines. In the original façade separation algorithm this additional information is not necessary but for vertical aerial images it is crucial for the outcome due to the fewer façade details caused by the lower resolution of the façade images. After this is done all façade strips are projected into the vertical aerial images and rectified. In a last step the façade separation is performed on the rectified façade strips.

3.1 Building Block Outlines

For each building block we have to determine its outline. The building objects obtained from the image classification are an approximation of the intersection of a façade with the ground. The goal is to isolate the contour of each building block. Initially, this contour is in the form of pixels in need of a vectorization. This has for a long time been studied and a choice of different methods exists. In our case we employ the recursive line simplification by Douglas-Peucker (1973). The goal is to replace the number of vertices in a piecewise linear curve. The contour pixels get replaced by straight lines, each defining one side of the building block. Figure 4 illustrates the result of this calculation for our test area.

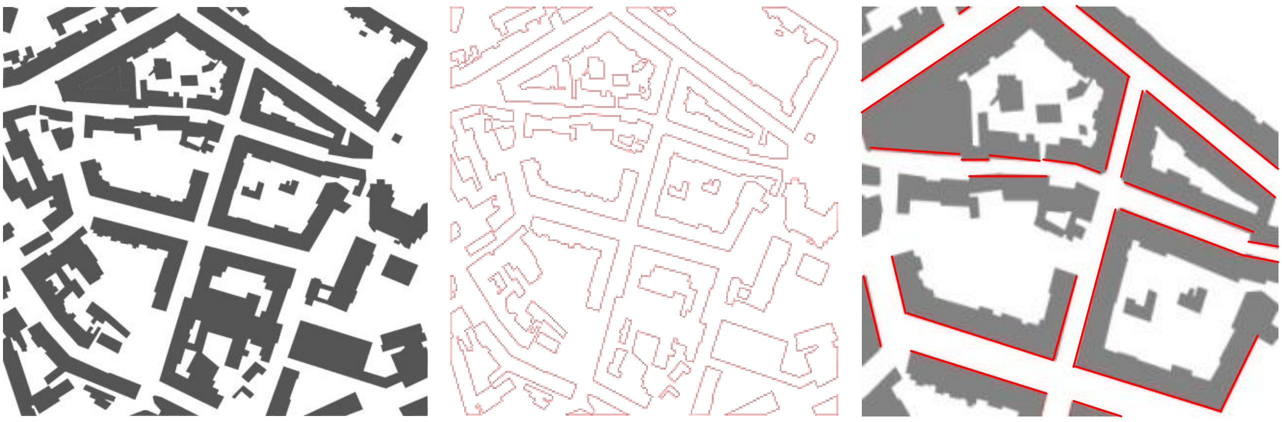


Figure 4. The classification layer "*building*" is based on color, texture and elevation values. (a) is the binary layer; (b) presents the contours as a raster image and (c) are the detected façade footprints highlighted in red

3.2 Incorporating Height Data

Because of the low resolution of the building façades this step is crucial for the outcome of this approach. The incorporation of the height values coming from the DSM has many benefits. First, we need it to add the third dimension to the 2D building block contours to determine the façade strip quadrilaterals. In the simplest case we assign one height value to the previously defined 2D lines. The result is a simple façade quadrilateral that can be projected into the aerial images (see Figure 6).

Second, using the elevation values with the 2D building contour allows us to split buildings with different heights already in advance. One may determine a measure of the building block symmetry for the elevations along the footprint: if façades get associated with different building heights, one may have reason to break the previously defined line into its parts. In our case these lines are split by calculating the gradient of the respective height profile (see Figure 5). If the gradient exceeds a certain threshold, in our case two meters, buildings are separated. The choice of an appropriate value for this threshold is crucial for the outcome to avoid false positives. These locations are stored and used as additional information for the following façade separation approach.

Third, the use of the height data allows us to segment façades in the vertical aerial images in more detail. Instead of four corner points for each façade strip we get a polygon that determines the outlines of the single building façades. This is helpful to enhance the separation results especially when the heights of attached buildings vary substantially (see Figure 6).

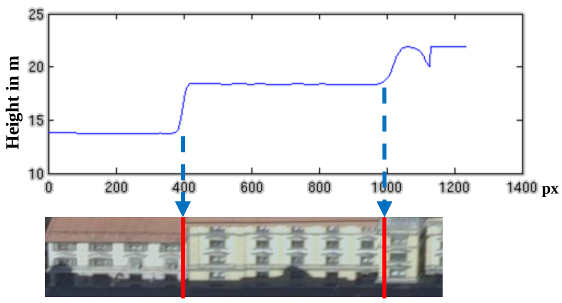


Figure 5. (Top) Height profile of a façade strip. (Bottom) Determined façade splits

After the generation of the façade strip quadrilaterals and the additional information coming from the elevation data we can project them into the aerial images and rectify them. Figure 6 illustrates a projected façade strip without additional height information and the enhanced version by incorporating different buildings heights.

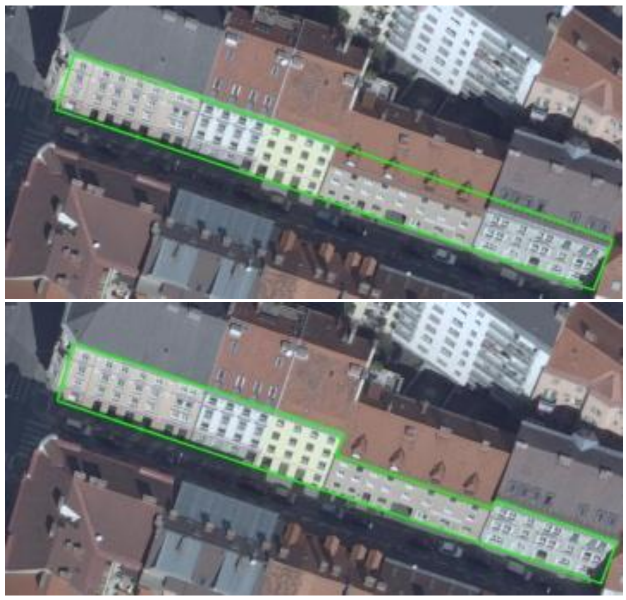


Figure 6. (Top) Projected façade quadrilaterals (Bottom) projected façade polygon

3.3 Façade Separation Algorithm

In this section we describe the façade separation algorithm in more detail. The result of the previous step produces façade areas, not individual façades per building. To be able to interpret the single buildings we have to split these façade areas into single façades. The applied algorithm was introduced by Wendel (2009) and consists of two major steps: First, repetitive patterns are detected in the façade images. In a second step the resulting pairs of interest points are then used to separate the façades.

3.3.1 Finding Repetitive Patterns

In a first step repetitive patterns in the images get associated with façades. The method uses Harris corners as interest points.

Figure 7 illustrates the detected Harris corners for one façade area (Harris et. al 1988). In a next step the color profile on a straight line between every interest point and the 30 nearest neighbors is calculated. The color profiles are constructed using a 20-dimensional normalized descriptor for each of the three colors RGB, in total thus with 60-dimensions. Finally the 60-dimensional descriptor is normalized. A *kd*-tree method is then used for matching the descriptors. Matches with more than 10 descriptors are eliminated already in advance because they are not discriminative enough. In a last step the repetitive patterns are located in a voting matrix. Within this voting step all matching profiles vote for the similarity of the respective pair of start- and endpoints. The results are stored in a list of contributing profiles for every possible pair. To strengthen the results for repetitive patterns on façades we eliminate all matches between two profiles that are too far apart or too close to each other.

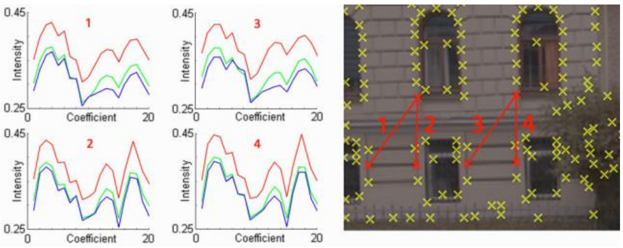


Figure 7: Detected Harris corners and the extracted intensity profiles consisting of 20 values for every RGB channel (taken from Wendel et al. 2010)

3.3.2 Façade Separation

In a next step the processing of the single façade is discussed in more detail. Due to the natural settings of objects in these images we assume that repetitive patterns occur along the horizontal direction and the separation of the façades occurs in the vertical direction. Therefore the lines between the matched interest points are projected into the horizontal axis constructing a match cost histogram. The match count is normalized to obtain the percentage of all matches, the repetition likelihood. Then the façades are segmented by determining a separation area (area where one façade ends and the next begins). This is done by defining areas with a low likelihood as separation areas and areas with high likelihood as repetitive areas. To be able to determine the exact split between two façades in a last step we look for the global maximum in these areas. An illustration of this step can be found in Figure 8.

4. EXPERIMENTAL RESULTS

The experiments were performed for a test area that covers 400m x 400m near the urban core of the city of Graz (Austria) with 186 different buildings. The vertical aerial photography was taken with a GSD of 10 cm and 80% forward and 60% sideward overlaps, using the large format digital aerial camera Microsoft UltraCam-X. These 186 buildings are grouped into 65 major building blocks with different sizes and shapes (see Figure 3b). Because some of these blocks consist of just one single building, in a preprocessing step all these blocks were eliminated. Furthermore just façade areas with a minimal length were used for this evaluation. We performed the evaluation of the proposed method using manually labeled ground truth for 31 façade images containing a total of 121 single façades using three different settings.

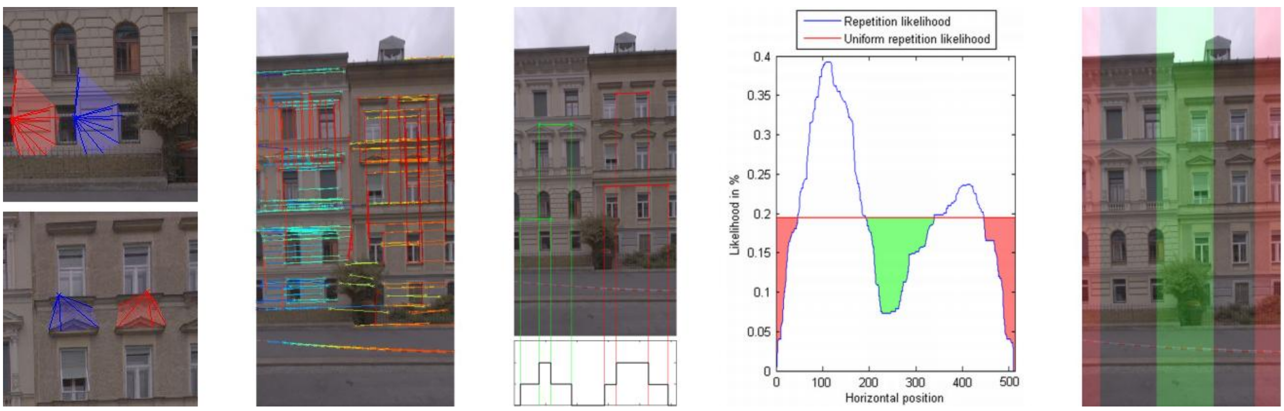
4.1 Façade Separation without Elevation Data

For our first experiments we have used only the vertical aerial images without additional height information to determine the impact of this information on the outcome of the processing. In this case we have used the RGB-information coming from the aerial imagery and one height value for the whole façade area (see Figure 6a). We performed the calculation for 31 façade images and achieved a success rate of 65 % that corresponds to 79 detected building façades out of 121. Figure 9a shows the result using these settings. Reason for the low detection rate are the low resolution of the aerial images that handicaps the extraction of repetitive patterns.

4.2 Façade Separation including Elevation Data

In our second experiment we have incorporated height information in our façade separation. This time also preliminary façade splits detected in Section 3.2 are included. Figure 5 illustrates an example of such a detected façade split. Like before, we have performed our experiments for 31 façades and achieve an increase of the success rate from 65% to 88%; this corresponds to 106 detected building façades. Figure 2 and Figure 9b show the results for façade separation including this additional information. As one can notice all façades could be separated correctly.

In our experiments we had 21 false positives that occur due to the nature of the used method (see Figure 9c).



(a) (b) (c) (d) (e)

Figure 8: From street side data to separation . (a) Matching of arbitrary areas (b) Detected repetitive patterns (color-coded lines) (c) Projection results in a match count along the horizontal axis (d) Thresholding the repetition likelihood with the uniform repetition likelihood (e) Resulting repetitive areas and separation areas (green) (taken from Wendel et al. 2010).

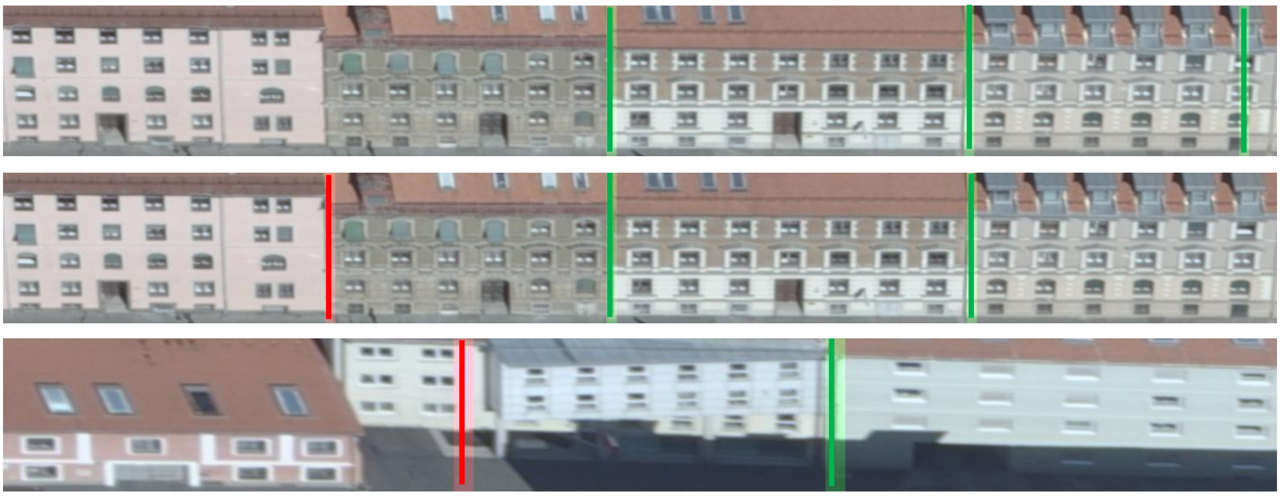


Figure 9: Result of façade separation, red and green lines are detected splits using height and repetitiveness, respectively. (top) without height information, (middle) including height information and (bottom) façade with a false positive (marked in red)

4.3 Façade Separation using Street Side Images

The question is: "*Are these results at 88% good or not?"* To be able to answer this question we consider façade detection in high resolution and highly overlapping street side images.

Tests with street side imagery have been performed in a subset of our test area with 9 separate building façades shown in 20-50 overlapping photos. The street-side images are taken in a forward look so that the façades are shown under an oblique angle. This helps in evaluating the influence of the perspective distortion. A detection rate was achieved of 97% (Wendel et al. 2010), whereby the facades were plane and best case.

We need to compare the 88% success from vertical aerial images with the 97% success in street side images. The difference is far less than expected when one considers that the vertical aerial imagery nominally has 10 cm pixels, but a very oblique viewing perspective of façades, whereas the street side imagery is in the 1 to 2 cm range and looks far less obliquely at façades. Therefore we conclude the façade separation from vertical aerial images to be feasible and successful.

5. CONCLUSION

Aerial photography is a work horse for urban mapping and exists for all urban spaces. It contains information about façades and roofs that needs to get extracted. We have presented a novel approach to separate building façades using vertical aerial images. Our approach can identify building footprints and uses additional elevation data for the façade separation. Initial work succeeds in finding relevant information with accuracies in the range of 88% and more.

We plan to evaluate further the performance of the proposed method by using more test data coming from different urban environments.

6. REFERENCES

Douglas D., Peucker T. 1973. Algorithms for the reduction of the number of points required to represent a digitized line or its caricature, *The Canadian Cartographer* pp. 112-122.

Feucht R. 2007. *„Flächenangaben im österreichischen Kataster"*, Diploma Thesis, Institute for Geoinformation and Cartography at Vienna University of Technology, 102 pages.

Harris C., Stephens M. 1988. A combined corner and edge detector. In: Proceedings of the Alvey Vision Conference. Volume 15. 50 pages.

Klaus A. 2007. *Object Reconstruction from Image Sequences*. Dissertation, Graz University of Technology.

Kluckner S., Bischof H. 2009. Semantic Classification by Covariance Descriptors within a Randomized Forest. *Proceedings of the IEEE International Conference on Computer Vision, Workshop on 3D Representation for Recognition* (3dRR-09)

Wendel A. 2009. *Façade Segmentation from Streetside Images.* Master's Thesis, Graz University of Technology, 72 pages.

Wendel A., Donoser M., Bischof H. 2010. Unsupervised façade segmentation using repetitive patterns. In: *Proceedings of the 32nd Annual Symposium of the German Association for Pattern Recognition* (DAGM'10), Springer LNCS 6376, pp. 51-60.

Zebedin L., Klaus A., Gruber-Geymayer B., Karner K. 2006. Towards 3D map generation from digital aerial images. *ISPRS Journal of Photogrammetry and Remote Sensing*, Volume 60, Issue 6, September 2006, Pages 413-427