Windows Detection Using K-means in CIE-Lab Color Space

Michal Recky, Franz Leberl
Institute for Computer Graphics and Vision, TU Graz
recky@icg.tugraz.at

Abstract

In this paper, we present a method for window detection, robust enough to process complex façades of historical buildings. This method is able to provide results even for facades under severe perspective distortion. Our algorithm is able to detect many different window types and does not require a learning step. We achieve these features thanks to an extended gradient projection method and introduction of a façade color descriptor based on a k-means clustering in a CIE-Lab color space into the process. This method is an important step towards creating large 3D city models in an automated workflow from large online image databases, or industrial systems. As such, it was designed to provide a high level of robustness for processing a large variety of façade types.

1. Introduction

Modelling large urban areas has become a challenging topic in recent years. The general goal in this research field is the recognition of important objects in the real urban scenes, creation of corresponding 3D model and the visualization of the model. Reconstruction of the buildings is considered a key part in this workflow [2]. The motivation for our work is to provide precise data for the building reconstruction and the need to interpret scenes as part of establishing an Internet-hosted Exabyte 3D World model [6]. The need to address the human scale of such a World model leads one to consider street side images, either via the use of an organized industrial sensor approach [3] or via crowd sourcing based on user-provided imagery [11]. In this paper, we present a method for the analysis of a building façade, identification of levels containing windows and the detection of windows.

We consider a single image scenario, where only one image of the examined building is available. This process can be easily extended into a multi-view scenario with improved accuracy of detection. The method described in this paper is designed to process complex facades of historical building, containing a large variety of ornaments, arches, patterns and divisions. The algorithm is able to process façade projected from a wide angle to the façade normal, therefore under a high perspective distortion. For the purpose of testing, we created a database of various historical and modern buildings located in the urban core of the Austrian city Graz and its peripherals. The images also exhibit a variety in different lighting and weather conditions.

2. Gradient projection façade analysis

In this section we will introduce an algorithm for processing a single façade located in a single image. Our work is based on horizontal/vertical gradient projection approaches, primary on a work of Lee, Nevatia [8]. This is a natural approach for the façade analysis, but in the original form, it is not suitable for complex facades, where the high levels of gradient in vertical/horizontal direction can be located also outside the windows area (see Figure 1). We therefore introduce a new – extended gradient method to deal with this problem. In our approach, the gradient projection is used to create a more detailed division of the façade. The resulting division is similar to the super-pixel based semantic segmentation method [5][9]. We use the k-means color clustering method [1] to perform labeling.

As an input, we consider the building facades to be identified in the single images by their borders. This can be achieved in several ways. In the approach of Lee, Nevatia [7] the aerial model of the scene with un-textured building frames is available and the rectified facades are obtained by projecting a digital image into this model. For some images, we don't have a frame model available. In these cases, the facades were labeled manually.

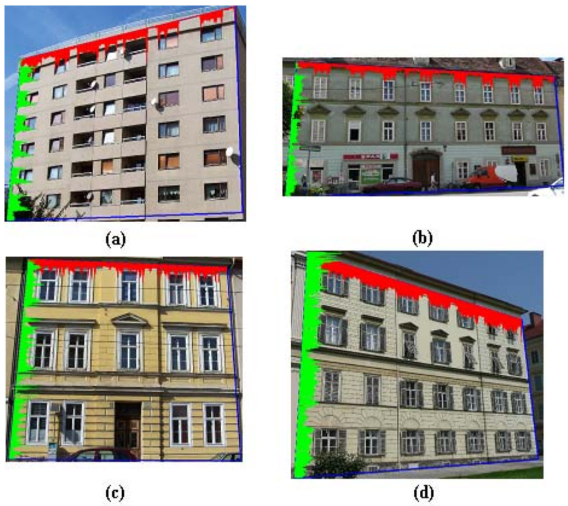

Figure 1: Four different types of facades in our database. The vertical gradient projection is marked in a green and the horizontal gradient projection is in a red color. Façade (a) is relative simple and the windows can be directly extracted from the projections. Façade (b) has additional horizontal structures, which make the horizontal separation difficult. In facades (c) and (d) both projections are highly non-regular and the extraction of windows is more complex.

2.1. Division of the façade into levels

In the first step, we compute the vertical projection of the gradient. In a simple façade, the strongest responses in the projection would be at the top and the bottom frames of the windows. When combined with the horizontal projection and the appropriate threshold, this information would be enough to estimate the borders of the windows. In the complex facades of the European city cores, there are usually other objects which present even stronger responses, than window frames. The most common are the rims, row of arches, brick-like patterns, color strips or columns. Similar problem can be observed for modern architecture facades with balconies, different color paintings or wires. For the facades containing these objects, the simple combination of the horizontal and vertical projection will not give us required results.

In our approach, we use the vertical projection to establish a horizontal division of the façade. For each local peak in the vertical projection a horizontal separator line is created. In this step, the façade is divided into a set of levels (bordered by separator lines).

In the next step, each level is processed separately. Firstly, the horizontal projection of gradient is computed for each level. Subsequently, each level is divided into a set of blocks. The application of threshold on the horizontal projection in each level will provide the borders for the block. The areas with the overall projected gradient above the threshold and the areas below the threshold are separated into different blocks.

3. Description of the façade blocks

Our next step is to decide if the block is part of the window, or part of the façade. This is done in an iterative process, where in each loop, the decision for each block is made, if it is part of façade, or not. The decisions are computed based on the size of the block, color and the gradient content of the block. The reason to use the color information in the process is for the identification of the areas between windows and the areas of the façade with high gradient content. Previous works on horizontal/vertical projection of building façade used mostly gradient information, which is usually suitable for identifying smooth non-window levels. For the complex facades of historical buildings, the gradient itself is not sufficient and the introduction of another descriptor is necessary.

Building facades usually consist of the large areas of uniform color, or small number of different colors (in clusters). The changes of illumination (mostly shadows) are often present and the same color may be displayed in different level of brightness. Standard area color descriptors, like RGB, or HSV histograms proved to be insufficient for this task, as they did not provide a correct description of a distribution of color areas. For this reason, we used the k-means clustering in CIE-Lab color space (see Figure 2). Selection of this color space can provide two significant advantages:

1. Euclidean distance of two colors in CIE-Lab space is directly proportional to the visual similarity of the colors. This can provide simple metric for a clustering.
2. The clustering can be performed only in "a", "b" space, which represent the color value component. The "L" component in CIE-Lab space represents the luminosity. A single value of L computed as a mean of each color in the cluster can be used as representative for each cluster to cope with shadows and illumination problems.

Figure 2: k-means clustering in CIE-Lab color space. Top row – façade with uniform color and it's clustering visualization on the right side. The façade in the bottom row consist of areas with several different colors. This non-uniformity express by itself as several different clusters in a color space.

In each iteration step, we are provided with blocks labeled as the façade. The color of pixels in these blocks is transformed into a CIE-Lab color space and the process of clustering is performed. The number of clusters is computed as follows:

1. k = 0; S = Ø
2. $\forall c \in \mathrm{F} : if\left(\neg\left(\forall s \in \mathrm{S} : \left|c - s\right| \leq th\right)\right) c \rightarrow \mathrm{S}$
3. k = |S|

Where k is the number of clusters, F is the set of façade pixel's colors and *th* is threshold for the distance of two colors belonging into the same cluster.

In the initial step, the blocks horizontally longer than 1/3 of the façade width are automatically labeled as façade blocks. In each subsequent step blocks are re-labeled according to the actual color description of the façade. When all blocks in one level are labeled as the façade, the entire level is excluded from the reclassification, but still contributes to the area color descriptor. After several iterations – usually less then five, depending on the complexity of the façade – there are no more changes in labeling. After this step, all blocks are labeled as façade, or unlabeled. Window blocks are identified as a non-façade blocks with gradient content.

4. Results

In our experiments, we use the dataset of 19 facades with 392 windows manually labeled. Of these, 369 windows were detected by the algorithm.

In the next step, we compare our method with the typical gradient projection method, as described in the paper of Lee, Nevatia [8]. Precision of window placement (in percentage) is computed as ratio between the width/height of detected window and the closest manually labeled window width/height to façade dimension. In this experiment, we examine a relationship between the gradient content, as the measure of façade complexity and the precision of window placement. Gradient content of the façade is computed as an average of gradient value {0,...,255} for each façade pixel (windows pixels are not considered as part of the façade in this case). The results are displayed in the Figure 3.

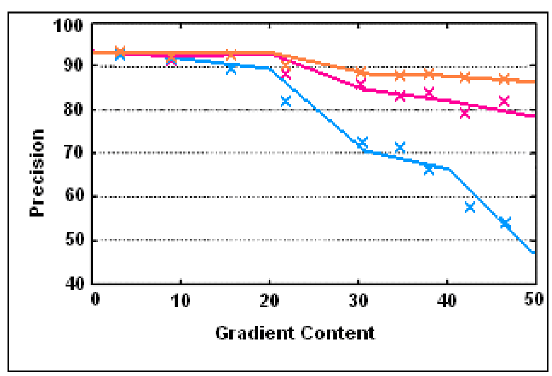

Figure 3: Relationship between the gradient content of the façade (excluding windows) and the precision of windows detection (in percentage). The blue line is displaying the relationship for the standard gradient projection method; the red line is for the extended gradient method, using color histograms; the orange line is for the extended gradient method using k-means in CIE-Lab color space. Crosses mark some plotted façade averages.

From the results of this experiment we can conclude, that method described in this paper (extended gradient method) performs significantly better for the facades with high gradient content. Most historical building in our database (city core in Graz) has a gradient content between 40 and 50. In this group, the precision of window placement can improve up to 40%, using k-means clustering. Some examples of buildings with high gradient content and the resulting window detection can be observed in Figure 4.

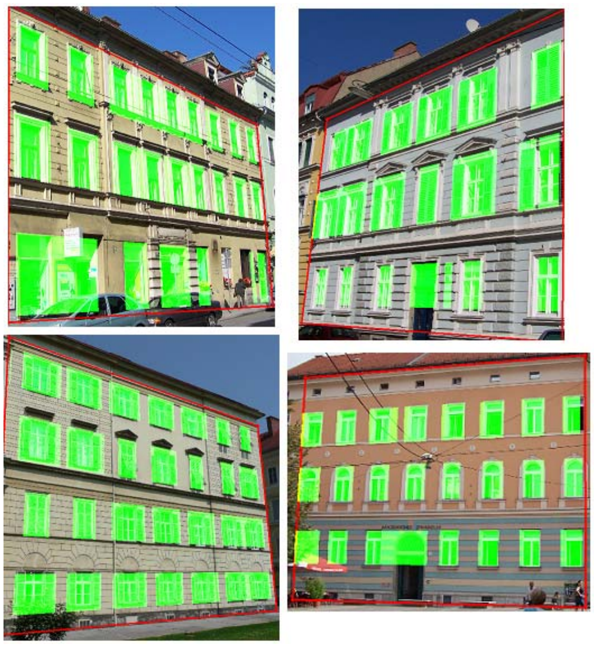

Figure 4: An example of windows projected back to the facades. All buildings have complex facades and are under perspective distortion.

5. Conclusion

In this paper, we presented a method for detection of windows in complex facades. The method is robust enough to process a large diversity of facades, from simple modern architecture, to the historical facades containing many decorative objects. This method can be considered as a step between the gradient projection approaches [10][8] and the general segmentation methods [4]. In addition to the gradient information, we used the CIE-Lab based color descriptor for labeling. This approach exploited the native uniform coloring of building facades and provided us with an advantage over previous methods.

Even through this paper is focused on a window detection, we observed that many typical façade objects (arches, rims, columns, rectangular patterns,…) have a specific signature in the gradient projections, and thanks to the division into multiple levels, they can be identified as well. Therefore in our future work, we will focus on the more general façade analysis.

Acknowledgments: This work has been supported by the Austrian Science Fund (FWF) under the doctoral program Confluence of Vision and Graphics W1209

References

[1] C.N. Anagnostopoulos, et.al., Tile Classification Using the CIELAB Color Model *Computational Science – ICCS* 2005

[2] J. Čech , R. Šára. Windowpane Detection based on Maximum Aposteriori Probability Labeling *Barneva, R. P. & Brimkov, V. (ed.) Image Analysis - From Theory to Applications*, International Workshop on Combinatorial Image Analysis (IWCIA'08), pp. 3-11 2008

[3] N. Haala, M. Peter, J. Kremer, G. Hunter, Mobile LiDAR Mapping for 3D Point Cloud Collection in Urban Areas - a Performance Test. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences,* Vol. XXXVII, ISPRS Congress, Beijing, China 2008

[4] F. Han, Z. Song-Chun, Automatic Single View Building Reconstruction by Integrating Segmentation. *Computer Vision and Pattern Recognition Workshop,* pp 53 – 60 CVPRW 2004

[5] D. Hoeim, A.A. Efros, M. Herbert, Geometric context from a single image. In *Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference*, volume 1, pp. 654–661 Vol. 1, 2005.

[6] F. Leberl, M. Gruber, 3d-Models of the Human Habitat for the Internet. *Proceedings of Visigrapp, pp 7-15*, Lisbon 2009

[7] S.C. Lee, S.K. Jung, R. Nevatia, Automatic Integration of Facade Textures into 3D Building Models with a Projective Geometry Based Line Clustering, *Computer Graphics Forum (Euro Graphics)*, pp.511-519 March 2002

[8] S.C. Lee, R. Nevatia, Extraction and integration of window in a 3d building model from ground view images. *In Computer Vision and Pattern Recognition,. CVPR. Proceedings of the 2004 IEEE Computer Society*, Vol.2, pages 112-120. 2004

[9] M. Recky, F. Leberl, Semantic Segmentation of Street-Side Images. *Proceedings of the Annual OAGM Workshop. Published by the Austrian Computer Society in OCG*, pp 271 – 282, April 2009

[10] K. Schindler, J. Bauer, A model-based method for building reconstruction. *In HLK 03: Proceedings of the IEEE International Workshop* 2003

[11] N. Snavely, S.M. Seitz, R. Szeliski, Photo tourism: Exploring photo collections in 3d. *ACM Transactions on Graphics (TOG)*. 2006