ADAPTIVE PARAMETER TUNING FOR MORPHOLOGICAL SEGMENTATION OF BUILDING FAÇADE IMAGES

Andrés Serna, Jorge Hernández and Beatriz Marcotegui

Mines ParisTech
CMM - Centre de morphologie mathématique
Mathématiques et Systèmes
35 rue St Honoré 77305-Fontainebleau-Cedex, France

ABSTRACT

In this paper, we describe an automatic method to segment street level building façade images. Our approach assumes that images are rectified, cropped and their elements are aligned in a pseudo-regular structure. It is based on the accumulation of directional color gradients, combined with morphological filters in order to deal with textured façades. We propose an automatic parametrization of three filters included in the process: opening filter of size n_{op}, alternate sequential filter (ASF) of size n, and H-minima filter with contrast threshold h. This automatic selection offers robustness to noise, image resolution changes, shadows and textures. Quantitative and qualitative results are reported on a public annotated database, validating the good performances of our approach.

Index Terms— Mathematical morphology, façade segmentation, window detection, urban modeling

1. INTRODUCTION

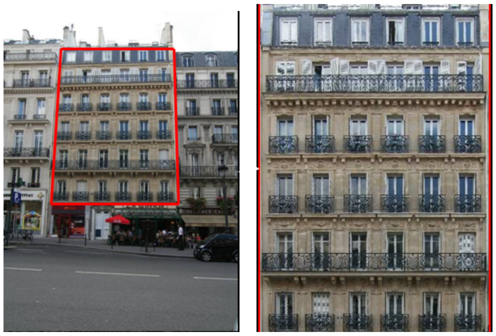
3D urban modeling is a very active research area. An increasing number of geographic applications, such as Google Earth, Microsoft Virtual Earth and Geoportail, are flourishing nowadays. These applications are not limited to 3D navigation, but also are considered as invaluable tools for environmental studies, urban planning, making accessibility diagnosis, etc. Some of these applications do not only require to look realistic, but also have to be faithful to reality.

Initially, virtual scenarios were created by infographic approaches, leading to time-consuming procedures, unsuitable for large-scale urban modeling. Procedural modeling allows to speed up the 3D virtual environment creation [1]. It is based on a set of rules, a grammar, defining a given architectural style. Procedural modeling approaches create realistic models in an efficient way, but a precise parametrization is required if the model has to be faithful to reality. Automatic analysis of façade images, combined with procedural modeling, allows to increase the productivity while remaining faithful to reality.

Many algorithms focusing on automatic façade analysis have been designed in the recent years. In general, existing methods use rectified and cropped images containing a single building. Usually, individual buildings are manually extracted. In [2], Lee and Nevatia develop a method based on thresholding of directional gradient projections. In [1], Müller et al. find repetitive architectural structures using mutual information to describe a single façade image. These methods are very sensitive to noise and fail if the building contains textured walls or balconies with different wrought iron designs, very common elements in Parisian Hausmannian architecture. In [3], Hernández et al. describe a method that automatically extracts an isolated building from a city block street level image. Besides, they extend Lee and Nevatia method introducing morphological filters in the directional gradient projections. These filters improve the robustness to textured façades. In [4], Teboul et al. learn a shape dictionary using random forest technique and publish an annotated database with 100 building images. In [5], Hammoudi extracts façade structures from 3D point clouds data using Hough transform. Finally, in [6], Pinte et al. combine color information and 3D point cloud data to improve the method robustness.

In this paper we focus on directional gradient projection techniques combined with morphological filters. We study the adaptive parameter tuning of these filters and evaluate the proposed algorithm on the cited public database. The paper is organized as follows. Section 2 describes the directional gradient projection technique, illustrated on a vertical splitting example. Section 3 describes the filter parametrization technique. Section 4 shows the performance of our method on Teboul's annotated database. Finally, Section 5 is devoted to conclude this work.

{andres.serna_morales, jorge.hernandez, beatriz.marcotegui} @mines-paristech.fr



(a) Street level image (b) Isolated building

Fig. 1. Automatic extraction of an isolated building [3].

2. SEGMENTATION OF BUILDING FAÇADES

The starting point for our approach is the method developed by Hernández et al. [3]. Therefore, it is assumed that input images are rectified and cropped, as shown in Fig. 1. Fig. 2 shows the diagram of the whole process and Fig. 3 illustrates intermediate images. First, a vertical gradient $G_y(x, y)$ detects horizontal contours (Fig. 3(a)). Then, a horizontal opening filter of size n_{op} is applied in order to eliminate the undesirable details. Fig. 3(c) shows the accumulation, column by column, of the vertical gradient. This 1D projected gradient contains peaks at window locations and valleys between them.

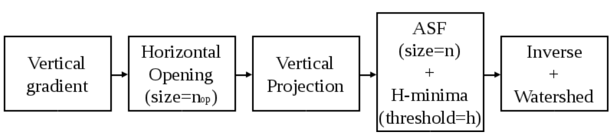
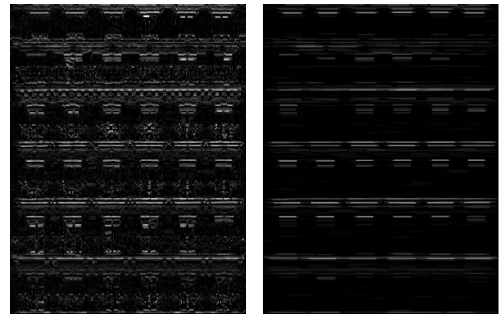


Fig. 2. Process scheme to compute vertical divisions.

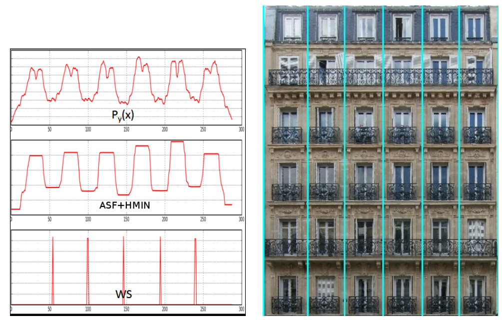
Afterwards, this projection is filtered in order to get a single maximum for each window. An Alternate Sequential Filter (ASF) of size n and a H-minima filter are used for this purpose. Finally, this profile is inverted and a watershed process computes the façade division. Fig. 3(d) shows the final result superimposed on the original image. Although a vertical splitting is shown, the same technique applies to horizontal splitting, just changing vertical by horizontal and vice versa.

3. FILTERING PARAMETRIZATION

The method introduced in the previous section leads to interesting results but relies on a good filter parametrization. Specifically, three parameters require tuning: the size n_{op} for the horizontal opening, the size n for the



(a) Vertical gradient: $G_y(x, y)$ (b) Horizontal Opening



(c) Vertical projection, filtering and watershed (d) Vertical splitting

Fig. 3. Vertical splitting

ASF, and the contrast threshold h for the H-minima filter. If these parameters are too small, the result will be over-segmented (Figures 5(a) and 5(b)). On the other hand, if they are too big, the result will be under-segmented (Fig. 5(d)). The aim of this section is to tune in an adaptive way these filter parameters, according to intrinsic image information. The parameter tuning of each step is explained below.

3.1. Opening filter parametrization

Windows are the image largest structures. Small details such as façade ornaments, wrought iron balconies or other noisy structures can produce fake divisions on projected profile $P_y(G_y)$. A morphological opening with a horizontal structuring element of size n_{op} is used in order to get rid of these details from gradient images. The selection of n_{op} is based on the pattern spectrum. Pattern spectrum plots the quantity of information filtered out by each opening γ_i: $(PS_i = \sum_{\forall pixel}(\gamma_{i-1} - \gamma_i))$. The resulting curve is also called *size distribution* because its peaks correspond to the prevailing sizes of the image structures [7].

Fig. 4 shows size distributions for different $G_y(x, y)$ images. These curves present an important peak for small size openings. This peak corresponds to noisy details. Note that this peak exists for the three images in spite of shadows (Fig. 11(b)), balconies (Fig. 10(b)), and vegetation (Fig. 11(a)). The opening size is chosen as the value i for which the pattern spectrum falls down under 25% of its maximum. This pattern spectrum analysis offers robustness to image resolution changes.

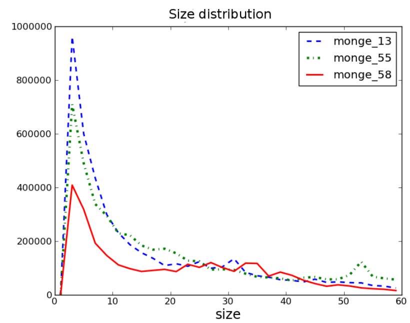


Fig. 4. Size distribution of $G_y(x, y)$ with a horizontal structuring element. Test images correspond to Figures 10 and 11

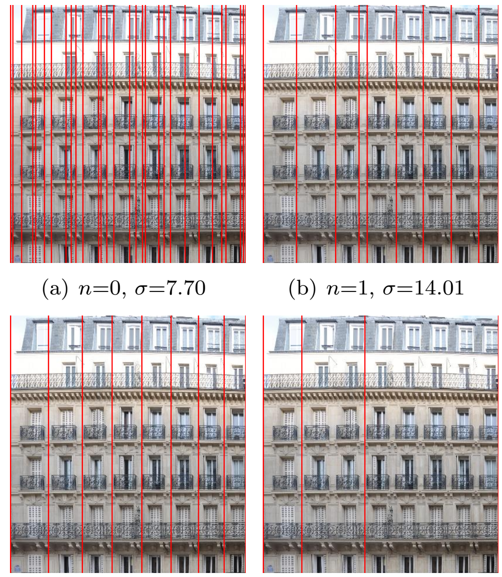
3.2. ASF filter parametrization

An ASF consists in a sequence of openings (γ) and closings (φ) of increasing sizes. The sequence starts with the filter of size 1 and ends with the filter of size n: $ASF_n(P_y(G_y)) = \gamma_n\varphi_n ... \gamma_2\varphi_2\gamma_1\varphi_1(P_y(G_y))$. This filter is particularly appropriated when the noise is present over a wide range of scales [8]. The filter size is chosen based on the faąde regularity. Several filters of different sizes are applied, and the one leading to the most regular result is chosen. The regularity is estimated by the standard deviation σ of the segmented façade division sizes. This filter applies to 1D profiles. Then, evaluating different sizes is not a time-consuming task. A frequency domain analysis of this profile would also be possible, but our approach is more robust to pseudo-periodic structures.

Fig. 5 shows the resulting vertical divisions for different filter sizes. Note that the filter size that minimizes the standard deviation, n=7, leads to a correct façade division.

3.3. H-minima filter parametrization

H-minima filter is a filtering tool based on a contrast criterion. More precisely, this transformation suppresses all minima whose contrast is lower than a given threshold h [9]. The contrast threshold h is chosen as a percentage of the dynamic of the extrema in the profile, that is $h \propto max(f) - min(f)$, where $f = ASF_n(P_y(G_y))$. This adaptive selection provides independence with respect to



(a) n=0, σ=7.70 (b) n=1, σ=14.01

(c) n=7, σ=6.24 (d) n=11, σ=68.90

Fig. 5. Façade divisions for different ASF sizes

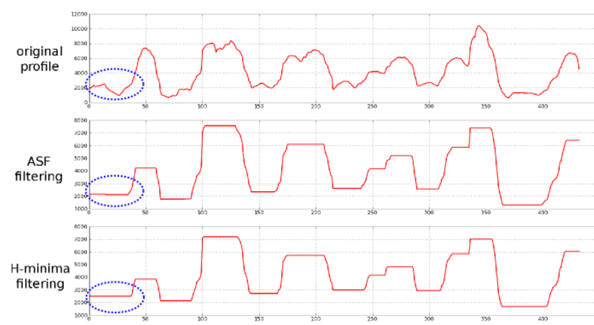


Fig. 6. ASF and H-minima filtering. The original profile corresponds to Fig. 10(a)

image resolution. Fig. 6 illustrates the effect of ASF and H-minima filtering. Note that the strongest filtering is carried out by the ASF, while the H-minima removes still remaining possible low contasted extrema, as shown in the left side of Fig. 6.

3.4. Window detection

We assume that there is only one column of windows per vertical division. Analyzing the extrema of the filtered profile $\tilde{P}_y(G_y)$, we found that minima pass through the wall while maxima pass through the windows. Using this information, we apply a constrained watershed on the projected horizontal gradient $P_y(G_x)$, taking the extrema of $\tilde{P}_y(G_y)$ as markers. Fig. 7 illustrates the process of window detection. Although we describe only the vertical case, the same approach applies to the horizontal detection.

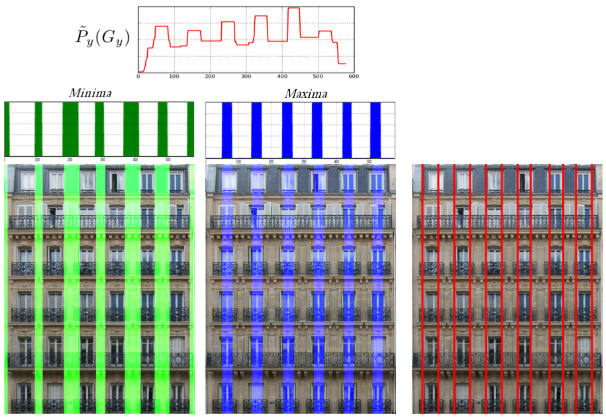


Fig. 7. Location of the vertical edge of the windows

4. EXPERIMENTS

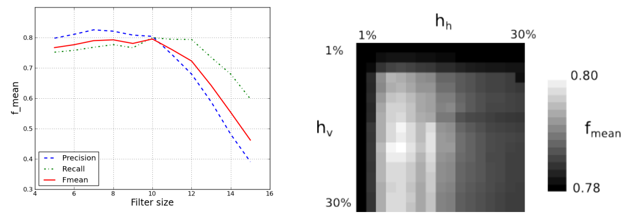
Our method is tested on the public database [4] that contains 100 annotated images. Images are rectified and various semantic elements are manually annotated. An example of this database is shown in Fig. 9. We evaluate our system on window localization with the classic $precision - recall - f_{mean}$ criteria. Note that our procedure detects windows including their corresponding balconies. In order to evaluate correctly the window detection performance, we remove from our detection the ground truth balconies regions

Fig. 8(a) shows the evaluation scores obtained with increasing ASF sizes. We can observe that the maximum f_{mean}=0.79 corresponds to filters of size between 7 and 10. If we use the ASF fitting method proposed in Section 3.2, we get the same score, the maximum in the figure, which proves the efficiency of the proposed tuning.

Once the parameters n_{op} and n are chosen according to the procedure aforementioned, we need to choose the best h threshold for the H-minima filter. Fig. 8(b) shows an exhaustive test varying h from 1% to 30% of the dynamic in the profiles. The best values found in the test correspond to h_v=14% and h_h=2.5% of the dynamic for the vertical and horizontal filter thresholds, respectively. Note that this parameter is not so critical since the lowest and highest f_{mean} correspond to 0.78 and 0.80, respectively. However, it improves the global performance up to 1% with respect to Fig. 8(a), where H-minima filter is not applied.

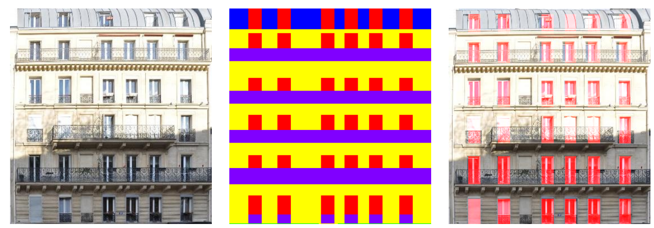
0

The results reported by Teboul et al. are $P = 0.58$, $R = 0.81$ and $f_{mean} = 0.68$ [4]. However, they only test 10 images of the database, while we have run our experiments on the whole dataset.



(a) ASF size (b) H-minima thresholds

Fig. 8. f_{mean} sensitivity to parameters n, h_h and h_v



(a) Original image (b) Ground truth (c) Detection

Fig. 9. Example of annotated image: (a) original image, (b) ground truth, and (c) our detection result

Using our proposed adaptive parameter tuning, the results are $P = 0.82$, $R = 0.79$ and $f_{mean} = 0.80$, which is much better than other results reported in the literature on the public Teboul's database. Qualitative results are shown in Figures 10 and 11. Fig. 10 shows examples in which the proposed method fails. Those images do not respect the regularity hypothesis on which our system is based: some window columns are almost adjacent. The standard deviation of division width is smaller when those columns are merged than when they are separated. Fig. 11 illustrates the robustness of our system to shadows, textures and images on which the distance between windows are pseudo-regular.

5. CONCLUSIONS

We propose an automatic parameter tuning of the three filters in the process: i) size n_{op} of the opening filter is deduced from the pattern spectrum analysis of gradient images. This filter removes texture details on the façade in order to avoid fake divisions. Moreover, its adaptive tuning offers robustness to image resolution changes. ii) Size n of the ASF is chosen as the value that minimizes the standard deviation σ of the segmented region sizes. This filter size has a strong influence on the result, as shown in Fig. 5. Its adaptive tuning according

(a) monge 34 (b) monge 58 (c) monge 63

Fig. 10. Examples of problematic results.

(a) monge 13 (b) monge 55 (c) monge 85

Fig. 11. Examples of correct results.

to the standard deviation of division sizes leads to the best result among all filter sizes. The adaptive result is just as good as the best result obtained with any filter size. This is because the whole dataset has the same resolution. If the database resolution were heterogeneous, the results would have been even better than the score with any filter parameter, because our parameter tuning would have been adapted to each image size. And, iii) contrast threshold h in the H-minima filter is chosen as a percentage of the dynamic of the extrema in the profile. This filter improves f_{mean} criterion of about 1% (from 79% to 80%). Its sensitivity in its whole dynamic range is very low. That means that the spurious maxima remaining after the ASF are very low contrasted, as shown in Fig. 6. In the horizontal direction, a very wide range leads to almost the same performances. On the other hand, in the vertical direction, filter thresholds between 5% and 15% of the dynamic score best.

The adaptive parameter tuning offers robustness to noise, image resolution changes, shadows and textures. These adaptive filters lead to the best performance score compared to any filter parameters tested in an exhaustive way. Thus, our approach is validated. Qualitative and quantitative results are reported. Our performances are much better than others reported in the literature on Teboul's public database.

In the future, the use of an adaptive opening operator, called ultimate opening, will be studied. This operator automatically adapts its size to the image structures, based on a contrast criterion.

6. ACKNOWLEDGEMENTS

The work reported in this paper has been performed as part of Cap Digital Business Cluster TerraNumerica project.

7. REFERENCES

[1] Pascal Müller, Gang Zeng, Peter Wonka, and Luc Van Gool, "Image-based procedural modeling of facades," *ACM Transactions on Graphics*, vol. 26, no. 3, pp. 85–93, 2007.

[2] Sung Chun Lee and Ram Nevatia, "Extraction and integration of window in a 3d building model from ground view images," *IEEE Computer Vision and Pattern Recognition*, vol. 02, pp. 113–120, 2004.

[3] Jorge Hernández and Beatriz Marcotegui, "Morphological segmentation of building façade images," in *IEEE International Conference on Image Processing, ICIP'09*, 2009, pp. 4029–4032.

[4] Olivier Teboul, Loïc Simon, Panagiotis Koutsourakis, and Nikos Paragios, "Segmentation of building facades using procedural shape priors," in *CVPR*, 2010, pp. 3105–3112.

[5] Karim Hammoudi, *Contributions to the 3D city modeling*, Ph.D. thesis, Université Paris-Est, 2011.

[6] Antoine Pinte, Caroline Baillard, and Eloïse Denis, "Modélisation de façades par analyse conjointe d'images terrestres et de données laser," *Revue française de photogrammétrie et de télédétection*, vol. 194, pp. 53–65, 2011.

[7] Georges Matheron, *Random Sets and Integral Geometry*, John Wiley & Sons, New York, 1975.

[8] Jean Serra, *Image Analysis and Mathematical Morphology*, vol. 2, Academic Press, London, 1988.

[9] Pierre Soille, *Morphological Image Analysis: Principles and Applications*, Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2003.