Graph Theory and Mean Shift Segmentation Based Classification of Building Facades

Beril Sirmacek

German Aerospace Center (DLR), Remote Sensing Technology Institute
PO Box 1116, 82234 Wessling, Germany
Email: Beril.Sirmacek@dlr.de

Abstract—Automatic classification of building facades has become a field of interest for researchers, since it can provide important solutions to many remote sensing research field. Urban scene understanding, mobile building recognition systems can be counted as the most important application fields. Classification results can be also used to increase the level of details in three-dimensional city models which are generated from aerial or satellite imagery. Characterization of stable regions in facades is also necessary for robust indexation and image retrieval.

Although in the literature many different facade classification techniques are proposed, because of the requirements of real-time applications robust and fast methods are still needed. Therefore in this study we propose a mean shift segmentation and graph theory based fast facade classification method. For this purpose, first mean shift segmentation is applied to color image and the connected pixels having similar color properties are labeled as one segment. Then, we benefit from graph theory to classify the segments as window, wall, and sky. The mass centers of the detected segments form the nodes of graph, and the neighborhood information is obtained through Delaunay triangulation. Next, a graph-cut is applied through graph of the image. Finally, we use a probabilistic verification step to compensate shortcomings of graph-cut step. Proposed algorithm is tested on a facade data set which is composed by many different kind of optical facade images which are taken from different places.

I. Introduction

Recently automatic classification of building facades has become a field of interest for researchers, since it can provide important solutions to many remote sensing research field like urban scene understanding and mobile building recognition. Detected doors and windows can be used in image registration when a mapping is applied to three-dimensional building model. Characterization of stable regions in facades is also necessary for robust image indexation and retrieval. Especially for mobile applications, very fast building understanding algorithms are still required.

In the last decade many different facade classification techniques are proposed. Pu and Vosselman [11] fused laser data and close-range images to reconstruct building facade details. They extracted windows and doors in both close-range optical and laser images by using Hough lines. Detected windows and doors helped them to register close-range optical and laser images. That shows the importance of the facade classification study in three-dimensional city modeling. Burochin et al. [3] proposed a segmentation method to detect repetitive structures like windows in close-range optical images. For segmentation they defined a model by considering shape and reflectance
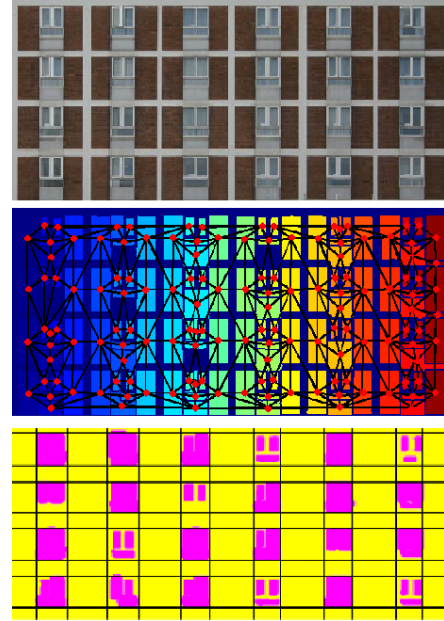

Fig. 1. Overview of the proposed approach ($Facade_1$ test image, segmentation result with facade graph, and classification result of the algorithm respectively.)

of a window, then they applied matching process to find correspondence between model and image. In [2], Ali et al. gave summary of the researches on window detection. They also proposed a window detection system based on cascade classifiers. In a following study, Ali et al. proposed a system to detect windows in laser scanner data. The laser distance information shows high variability in windows region. Therefore, they use these variations to detect windows [1]. Lee and Nevatia [8] proposed a robust system to detect windows in optical images. They extracted window boundaries searching for structures that satisfy regularity and symmetry rules. In addition to that, they extract three-dimensional models of windows by searching for image features. Teboul et al. [13] used shape grammars towards fixed tree representations which are able to capture a wide variety of building topologies for detailed facade segmentation. They obtained very high performance even for buildings which are partially occluded or which appears under different illumination conditions.

Most of the previous algorithms are computationally expensive and not suitable for real time applications. Because of the requirements of real-time applications robust and fast classification of facades is still an open research topic. In this study, we propose a segmentation and graph theory based fast facade classification method. Graph theory is widely used in computer vision applications for robust detection of objects. Sirmacek and Unsalan used graph theory to detect buildings automatically from satellite and aerial images. For this purpose, they have extracted SIFT features from the test image, and they used graph theory to represent spatial relations of extracted features. This representation helped them to find subgraphs which indicate building locations [12]. In another study, Zu used Delaunay triangulation for car detection in auto roads [16]. For this purpose, they extracted Canny edges of the image, and showed their relations with a graph which is obtained by Delaunay triangulation. They detected the car, by extracting the subgraph which satisfies predefined hypothesis about texture, edge, and corner properties. Herein, we also benefit from Delaunay triangulation because of its computational efficiency. First, we start with applying mean shift segmentation to optical image. Unfortunately, this segmentation result do not indicate windows and doors of the building directly. Then, we form a facade graph where the mass centers of segments form graph nodes. Graph edges are obtained by applying Delaunay triangulation to graph nodes. Next, we apply a graph-cut trough facade graph to obtain subgraphs which indicate windows. Finally, we make a probabilistic verification to compensate shortcomings of graph-cut step. In Fig. 1, overview of the proposed algorithm is demonstrated using $Facade_1$ test image from our database. The first image is the original orthorectified facade image. Rectangular structures between windows make window and door detection difficult in this image. The second image shows the segments which are obtained by mean shift segmentation. Constructed facade graph is also represented on this image. The third image, shows the classification result of the proposed method. Next, we explain facade graph construction.

II. GRAPH-THEORETIC FACADE MODEL

Graph theory is widely used in computer vision applications for robust detection of objects. Here we also cast the problem in terms of graph theory. We represent given facade image as a weighted graph. For this purpose, first we apply mean shift segmentation to the image as Comanicu et al. proposed [4]. The segmentation result is a matrix like $S(x, y)$ which holds each segment labeled by a different number. We provide mean shift segmentation result of our $Facade_1$ test image in the second row of Fig. 1. For better visualization, we represented each different segment with a different color. As can be seen in this result, mean shift segmentation reduces complexity of the problem, but the result can not directly indicate window locations. Therefore, we reconstruct a weighted facade graph.

A graph G is formed as $G = (N, E)$, where N holds the nodes of this graph, and E is the edge matrix showing the relations between these nodes. In our study, N holds the mass centers of the segments which are detected by mean shift algorithm. To reconstruct graph edges, we benefit from Delaunay triangulation, since only neighbor segments can form parts of a window. Using Delaunay triangulation, we connect only neighbor segments which also reduces complexity of graph. Therefore, E is a $M \times M$ matrix, where M represents the total number of segments in $S(x, y)$ matrix. E is defined as $E(i, j) = 1$ where $i, j \in \{1, 2, , ..., M\}$, if i and j nodes are connected by Delaunay triangulation. Otherwise, $E(i, j) = 0$ which means there is no edge between ith and jth nodes.

We also assign two weights to each graph edge. For ith and jth graph nodes if $E(i, j) = 1$, we assign two weights to this graph edge as d_{ij}, and w_{ij}. Here d_{ij} is the Euclidean distance between ith and jth nodes and defined as $d_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$. Second weight w_{ij} is the color distance between two segments, since color information also gives important cues for detection. In many computer vision applications, the CIELab color space is used since it mimics the human visual system [5]. CIELab color space bands enhance different colors best and minimize color variances. After transforming the RGB image into CIELab color space, we again obtain three bands as L, a, and b [10]. Here, L band corresponds to intensity of the image pixels. a, b bands contain chroma features of the image. They give information about the color independent of illumination. In the literature, many researchers used Euclidean distances of L, a, and b bands of images to find similar regions generally for segmentation purposes [7]. Here, we use a similar methodology to assign weights to the graph edges. We apply CIELab color transform to the original RGB image, and store mean of L, a, b value for each segment. For each segment a color feature is obtained as $[L_s^i, a_s^i, b_s^i]$ where $i \in \{1, ..., M\}$. Then, w_{ij} is defined as,

$$w_{ij} = \sqrt{(L_s^i - L_s^j)^2 + (a_s^i - a_s^j)^2 + (b_s^i + b_s^j)^2} \quad (1)$$

Constructed facade graph is represented in the second row of Fig. 1. Here, red circular shapes indicate locations of graph nodes, and black connections are the constructed graph edges. Next, we apply graph-cut to facade graph in order to detect windows.

III. GRAPH-CUT PROCESS TO DETECT WINDOWS AND DOORS

We obtained facade graph G and two weights for each edge of the graph. Here, our purpose is to use this graph to detect windows. Therefore, we cut some edges of G based on color and distance criterias. We assume that parts of the same window, or parts of the same wall should be close to each other, and should also be in similar colors. Therefore, to locate windows, we cut edges between nodes having different colors and having large Euclidean distance between them. We form a new subgraphs as $G^s = (N^s, E^s)$, where E^s is defined as below.

$$E^s(i, j) = \begin{cases} 1 & \text{if } (E(i, j) = 1) \wedge (d_{ij} < \epsilon_1) \wedge (w_{ij} < \epsilon_2) \\ 0 & \text{otherwise} \end{cases}$$

Here, ϵ_1 and ϵ_2 are the distance and color thresholds respectively to decide to cut a graph edge. To calculate ϵ_1 we use height of the image h. We assume h to be equal to the building height. Considering a four stage building, we assume ϵ_1 as $h/4$. Parts of the same window are always very close to each other, and they always satisfy this condition. For the parts of the same window, color features should be very similar and their difference should be close to zero. However, to give a tolerance we assume ϵ_2 as 5 in our study. In Experiments section, we will represent effects of ϵ_1 and ϵ_2 threshold values to the classification result.

After finding G^s subgraphs, we apply a last test to each E^s edge. If $E^s(i,j) = 1$, we apply compactness test to see if ith and jth segments are forming a compact shape or not. For this purpose we take union of two segments and assume them as one single block. Compactness is defined as the ratio of a region's area to the square of its perimeter, normalized by 4π [14]. If $compactness < 4\pi/60$ for this single block, we label this block as a window or one part of a window. Next, we apply last verification step to provide facade classification result.

IV. PROBABILISTIC VERIFICATION

Windows and doors of a building facade are generally in horizontal and vertical order. We use this property for verification step. We make a last probabilities test to eliminate false detected window and door segments. With this step, we also want to find missing doors and windows that could not be found with graph-cut process. For this purpose, we follow a probabilistic approach. For each window and door, we give a probabilistic vote $P_k(x,y)$, where $k \in \{1,2,...,K\}$ indicates the label of window. In Fig. 2, in the first and second image we give a sample window segment, and its probability matrix $P_k(x,y)$ respectively. After each window gives a probability to building facade, we obtain probability matrix as $P(x,y) = \sum_{k=1}^{K} P_k(x,y)$ by summing probabilities given by each window and door segment. We provide obtained probability matrix for $Facade_1$ test image in the last column of Fig. 2. The darker area shows the regions with 0 probability. Therefore, there should not be any window segment in these regions. If there is a window segment in the regions where $P(x,y) = 0$, we eliminate this window segment. Then, we make a last control to find missing window segments. To this end, we threshold $P(x,y)$ probability matrix with Otsu's automatic thresholding method [9], and after thresholding we obtain $B(x,y)$ binary mask. If there is a segment in $B(x,y) \cup S(x,y)$ which also satisfies the compactness test, then we label this segment as a window. Detected windows can be seen in the last column of Fig. 1.

V. LABELING BUILDING WALL AND SKY

After labeling window and door segments, we assume that rest of the segments should be from wall or sky class. For detecting sky, we generated a training set by cutting sky from one of the images manually. We generated $[L^t, a^t, b^t]$ color feature by calculating mean of L, a, b values in training set.

Fig. 2. Sample window segment, probability matrix for the sample window segment, $P(x,y)$ probability matrix are represented from left to right respectively.

A segment having $[L_s^i, a_s^i, b_s^i]$ $i \in \{1,2,3,...,M\}$ is labeled as sky if $\sqrt{(L^t - L_s^i)^2 + (a^t - a_s^i)^2 + (b^t - b_s^i)^2} < \epsilon_2$. A segment which is not satisfying this condition is labeled as wall. A sample detected sky can be seen in Fig. 3.d. In literature more advanced algorithms are proposed to detect sky robustly from color images [15], [6]. We will also try to improve our sky detection method on a larger data set in our future studies.

VI. EXPERIMENTS

We tested the proposed facade classification algorithm on 15 different facade images. We provide some of the experimental results in Fig. 3. Original images and the classification results can be seen in the left and right column respectively. We compared classification results with manually labeled building facades using following equation,

$$TP = 100 \times \frac{\sum_{x=1}^{X} \sum_{y=1}^{Y} (B_w(x,y) \cup B_{gth}(x,y))}{\sum_{x=1}^{X} \sum_{y=1}^{Y} B_{gth}(x,y)} \quad (2)$$

$$FA = 100 \times \frac{\sum_{x=1}^{X} \sum_{y=1}^{Y} [B_w(x,y) - (B_w(x,y) \cup B_{gth}(x,y))]}{\sum_{x=1}^{X} \sum_{y=1}^{Y} B_{gth}(x,y)} \quad (3)$$

where TP indicates percentage of true detected window areas. $B_w(x,y)$ and $B_{gth}(x,y)$ indicates binary mask of detected windows segments, and binary groundtruth mask respectively. In our test data set, for 15 facade images TP and FA are calculated as 86.65% and 30.16% respectively.

In Table I and II, we represent window detection performances for different ϵ_1 and ϵ_2 values. As can be seen in Table I, selection of ϵ_1 value do not make a significant change on performance. However, different ϵ_2 values may affect FA slightly, since when ϵ_2 color threshold is increasing, different colored wall segments may be merged to window segments. Computation time of the algorithm is impressive. For $Facade_1$ image which has 24 windows, computation time is only 8.9 seconds using a Core2Quad 2.66GHz PC and Matlab coding environment.

VII. CONCLUSIONS AND FUTURE WORK

Herein, we present a graph theoretic method for fast facade image classification. We use mean shift method for segmenting building facade image. Then, we benefit from graph theory to detect windows and doors. For this purpose, we generate a facade graph where mass centers of detected segments form the nodes of the graph. Edges of this graph are constructed
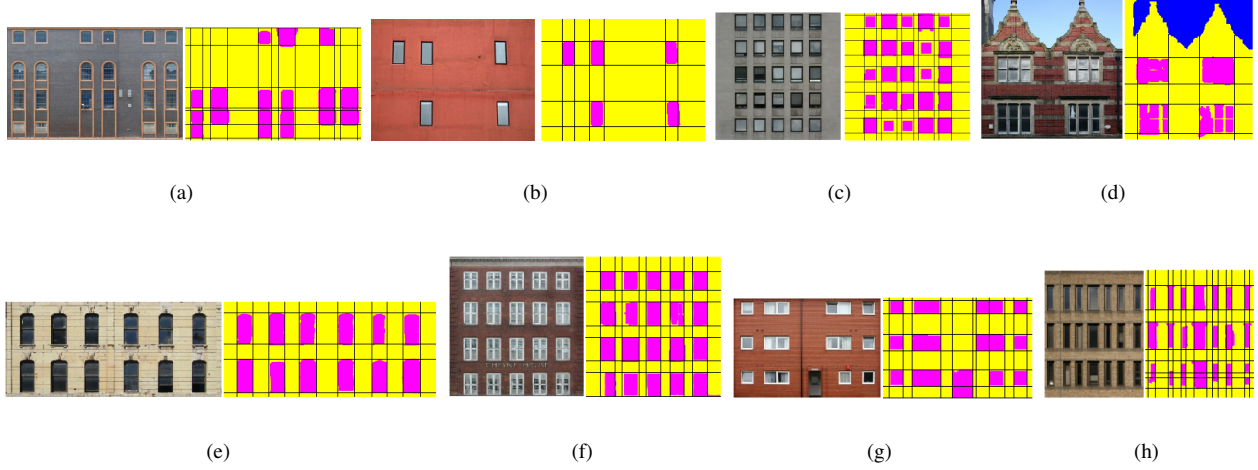
(a) (b) (c) (d)

(e) (f) (g) (h)

Fig. 3. Experimental results for $Facade_{2-9}$ test images from our data set.

ϵ_1	TP (%)	FA (%)
$h/2$	85.43	7.80
$h/3$	85.57	8.94
$h/4$	85.57	8.94
$h/5$	85.57	8.94
$h/6$	86.15	9.94

ϵ_2	TP (%)	FA (%)
2	71.50	8.05
3	81.46	8.58
4	85.15	8.79
5	85.57	8.94
6	87.64	9.13
7	87.64	9.13
25	87.93	11.18

using Delaunay triangulation method. Windows and doors are detected with applying a graph-cut process to facade graph. We used a probabilistic verification step to eliminate shortcomings of graph-cut step, and also to detect missing window and door segments. We also label the sky by checking the color properties of the segments.

Obtained experimental results demonstrate the potential of the method for real time applications like mobile building recognition, or image registration to three-dimensional building models. Still these are the first results. We are working on optimal selection of given parameters, and increasing number of classes.

REFERENCES

[1] H. Ali, B. Ahmed, and G. Paar, "Robust window detection from 3d laser scanner data," *Congress on Image and Signal Processing (CISP'08)*, vol. 2, pp. 115–118, 2008.

[2] H. Ali, C. Seifert, N. Jindal, L. Paletta, and G. Paar, "Window detection in facades," *14th International Conference on Image Analysis and Processing (ICIAP)*, vol. 1, pp. 837–842, 2007.

[3] J. Burochin, O. Tournaire, and P. Nicolas, "An unsupervised hierarchical segmentation of a facade building image in elementary 2d models," *ISPRS Workshop on Object Extraction for 3D City Models, Road Databases and Traffic Monitoring - Concepts, Algorithms and Evaluation (CMRT'09)*, 2009.

[4] D. Comanicu and P. Meer, "Mean shift: A robust approach toward feature space analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, pp. 603–619, May 2002.

[5] M. Fairchild, "Color appearance models," *Addison-Wesley*, 1998.

[6] A. Gallagher, J. Luo, and W. Hao, "Improved blue sky detection using polynomial model fit," *International Conference on Image Processing (ICIP'04)*, vol. 4, pp. 2367–2370, 2004.

[7] L. Haoting, G. Jiang, and L. Wang, "Multiple objects tracking based on snake model and selective attention mechanism," *International Journal of Information Technology*, vol. 12 (2), pp. 76–86, 2006.

[8] S. Lee and R. Nevatia, "Extraction and integration if window in a 3d building model from ground view images," *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2, no. 1, pp. 113–120, 2004.

[9] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Transactions on System, Man, and Cybernetics*, vol. 9 (1), pp. 62–66, 2009.

[10] G. Paschos, "Perceptually uniform color spaces for color texture analysis: an empirical evaluation," *IEEE Transactions on Image Processing*, vol. 10, pp. 932–937, 2001.

[11] W. Pu and G. Vosselman, "Refining building facade models with images," *ISPRS Workshop on Object Extraction for 3D City Models, Road Databases and Traffic Monitoring - Concepts, Algorithms and Evaluation (CMRT'09)*, 2009.

[12] B. Sirmacek and C. Unsalan, "Urban area and building detection using sift keypoints and graph theory," *IEEE Transactions in Geoscience and Remote Sensing*, vol. 47 (4), pp. 1156–1167, 2009.

[13] O. Teboul, L. Simon, P. Koutsourakis, and N. Paragios, "Segmentation of building facades using procedural shape priors," *Computer Vision and Pattern Recognition, CVPR'10*, vol. 1, pp. 3105–3112, 2010.

[14] C. Unsalan and K. Boyer, "A system to detect houses and residential street networks in multispectral satellite images," *Computer Vision and Image Understanding*, vol. 98, pp. 432–461, 2005.

[15] B. Zafarifar and P. H. N. de With, "Blue sky detection for picture quality enhancement," *8th International Conference on Advanced Concepts for Intelligent Vision Systems (ACIVS)*, vol. 4179, pp. 522–532, September, 2006.

[16] K. Zu, "Realtime obstacle detection and tracking based constrained delaunay triangulation," *IEEE Intelligent Transportation Systems Conference, ITSC'06*, pp. 548–553, 2006.