Architectural Style Classification of Building Facade Windows

Gayane Shalunts*[1], Yll Haxhimusa[2], and Robert Sablatnig[1]

[1] Vienna University of Technology
Institute of Computer Aided Automation
Computer Vision Lab
{shal,sab}@caa.tuwien.ac.at
[2] Vienna University of Technology
Institute of Computer Graphics and Algorithms
Pattern Recongition and Image Processing Lab
yll@prip.tuwien.ac.at

Abstract. Building facade classification by architectural styles allows categorization of large databases of building images into semantic categories belonging to certain historic periods, regions and cultural influences. Image databases sorted by architectural styles permit effective and fast image search for the purposes of content-based image retrieval, 3D reconstruction, 3D city-modeling, virtual tourism and indexing of cultural heritage buildings. Building facade classification is viewed as a task of classifying separate architectural structural elements, like windows, domes, towers, columns, etc, as every architectural style applies certain rules and characteristic forms for the design and construction of the structural parts mentioned. In the context of building facade architectural style classification the current paper objective is to classify the architectural style of facade windows. Typical windows belonging to Romanesque, Gothic and Renaissance/Baroque European main architectural periods are classified. The approach is based on clustering and learning of local features, applying intelligence that architects use to classify windows of the mentioned architectural styles in the training stage.

1 Introduction

Architectural styles are phases of development that classify architecture in the sense of historic periods, regions and cultural influences. Each architectural style defines certain forms, design rules, techniques and materials for building construction. As architectural styles developed from one another, they contain similar elements or modifications of the elements from the earlier periods. An automatic system for classification of building facade images by architectural styles will allow indexing of building databases into categories belonging to certain historic periods. This kind of a semantic categorization limits the search of building

* Supported by the Doctoral College on Computational Perception.

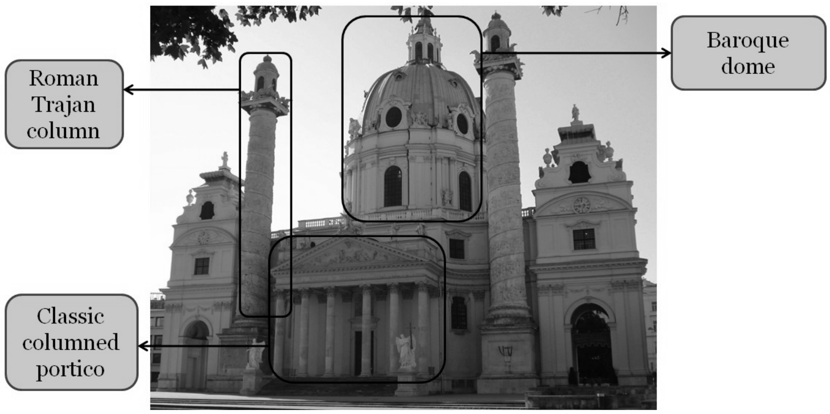


Fig. 1. Different architectural styles in St. Charles's Church in Vienna.

image databases to certain category portions for the purposes of building recognition [1, 2], Content Based Image Retrieval (CBIR) [3], 3D reconstruction, 3D city-modeling [4] and virtual tourism [5]. Architectural style classification system may also find its application in tourism, if provided with smart phones.

To the best knowledge of the authors there is no automatic system for classification of building facade images by architectural styles. Building facade images from online image databases either do not have any labels related architectural styles or such labels are inaccurate. If the observer does not have the knowledge how to classify architectural styles, he/she should search for the name of the image building and thus find out the architectural style of the mentioned building. If the building image does not have any annotations, it is impossible for the observer to find out to which architectural style the building belongs to. An automatic system for classification of architectural styles will solve this task.

Architectural style classification of the whole building is viewed as a voting mechanism of separate architectural elements, such as windows, domes, towers, columns, etc. This approach allows facade architectural style classification by a single structuring element, for example a window, in case of partly occluded facades. It is also appropriate for facades which are a mixture of architectural styles. In case of voting for different architectural styles by different architectural elements, the more important architectural elements are given heavier weights while voting. A typical example of a building designed in different architectural styles is St. Charles's Church in Vienna (Fig. 1), which includes Roman columns, a Classic columned portico and a Baroque dome. In this case the dome should be given a heavier weight than the columns and portico, as St. Charles's Church is considered a Baroque church.

In the scope of facade architectural style classification task by a voting mechanism of structural elements, the current paper focuses on classification of typical facade windows of the main European proceeding architectural styles:

a) Single arch b) Double arch c) Triple arch

Fig. 2. Romanesque windows.

– Romanesque (1050 A.D. - 1200 A.D.)
– Gothic (1150 A.D. - 1500 A.D.)
– Renaissance (1420 A.D. - 1550 A.D.)
– Baroque (1550 A.D. - 1750 A.D.)

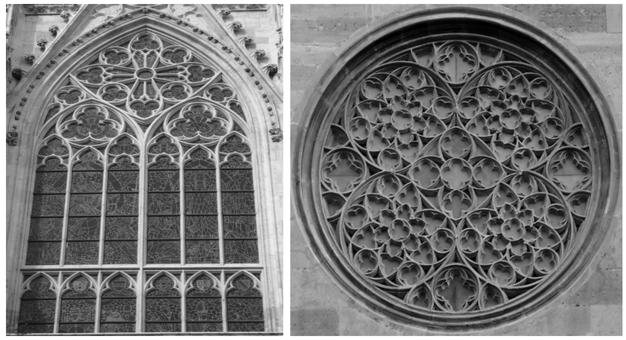
As there are methods like [6–8] for detection of windows on building facades, the current paper operates on an image database of bounding boxes of windows. Our approach is based on the fact that each architectural style applies certain geometrical rules for style typical window construction. This means that certain gradient directions are dominating in each window class. The methodology is based on clustering and learning of the local features to find out the image dominant gradient directions and thus categorize the classes of different architectural styles. Our system yields a classification rate of 95.16% while categorizing 3 architectural styles and 8 intra-class types.

The paper is organized as follows: Section 2 shows typical windows of Romanesque, Gothic, Renaissance/Baroque architectural styles which are classified. Section 3 explains the chosen method for the classification of the mentioned window types. The experiments and results of the classification are presented in Section 4. And finally Section 5 concludes the paper.

2 Typical Windows of the Classified Architectural Styles

For architectural style classification of windows typical window examples of Romanesque, Gothic, Renaissance/Baroque architectural periods are chosen. The characteristic feature of Romanesque windows is the single, double or triple round arch (Fig. 2a, b and c respectively), while Gothic style is very distinct with pointed arches (Fig. 3a) and rose windows (Fig. 3b).

For Baroque style window decorations like triangular and segmental pediments (Fig. 4a and b respectively) and balustrades (Fig. 4c) are characteristic.



a) Gothic pointed arch b) Gothic rose

Fig. 3. Gothic windows.

As Baroque evolved from Renaissance windows with triangular, segmental pediments and balustrades are also present on Renaissance buildings. In the case of the mentioned window types other features should be taken into account to differ between Baroque and Renaissance styles. Such features may be depth information, as Renaissance is considered 'planar classicism' and Baroque as 'sculpted classicism' or the analysis of the whole building facade structure. Our method overall catigorizes 3 window classes:

- Romanesque
- Gothic
- Baroque

and 8 intra-class types - Romanesque single, double and triple round arch windows, Gothic pointed arch and rose windows, Baroque windows with triangular, segmental pediments and balustrades. We classify between the 3 stated architectural classes, but not the 8 intra-class types, as our objective is architectural style classification.

In the scope of architectural style classification task it should be mentioned about architectural revivalism, which is a phenomenon of imitation of past architectural styles. The singularity of 19th century revivalism, as compared with earlier revivals, was that it revived several kinds of architecture at the same time [9]. These revived styles are also referred to as neo-styles, e.g. Gothic Revival is also referred to as neo-Gothic. Our approach does not differ between original and revival architectural styles, as only visual information is not enough for such a discrimination. Additional information related building date, location and materials is needed to differ between original and revival architectural styles.

a) Triangular pediment b) Segmental pediment c) Balustrade

Fig. 4. Baroque windows.

3 Bag of Words for Facade Window Classification

The task of classification of windows by architectural styles is highly complex, because of the high intra-class diversity as well as reflections present in window images. One can use different texture features [10, 11] as well as established shape descriptors [12, 13]. In this work we use a local feature-based approach, since it incorporates texture and gradients into an image descriptor. It is shown in [14] that shapes can be represented by local features (peaks and ridges). Since on window shapes of each class certain gradient directions are dominating, we use local features to describe shapes. One can use different local features, like Harris-Laplacian corner detectors [15, 16], difference of Gaussians corner detectors [17] or detectors based on regions [18, 19] and local image descriptors [17–19]. The goal is to extract characteristic gradient directions, like those describing pointed arch or triangular pediment (Fig. 2,3,4) and to minimize the influence of non-relevant features, like those from reflections and curtains.

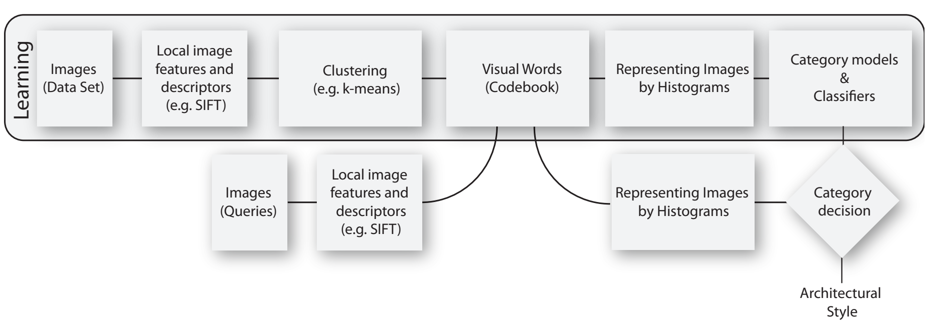


Fig. 5. Learning visual words and classification scheme.

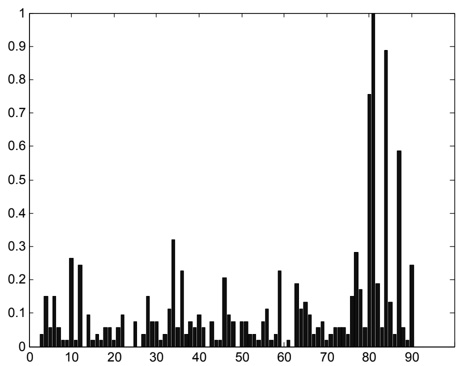
Classifying windows will be preceded by a method to classify facades, thus we choose the standard bag of words approach presented by Csurka et al [20] (Fig. 5). In the learning phase the Scale Invariant Feature Transform(SIFT) [17] is used to extract the information of gradient directions. After performing the difference of Gaussians on different octaves and finding minimas/maximas, i.e. finding interest points, we only perform rejection of interest points with low contrast by setting a low threshold. All interest points that lie on window edges are kept. Note that we do not follow the original work [17] in this step, i.e. we do not suppress the response of the filter along the edges. After finding the interest points we proceed to finding local image descriptors (SIFT image descriptors) and normalizing them. The number of local features is large, thus we use clustering to learn a visual vocabulary (codebook). The codebook of separate classes is made by searching for the visual cluster centers using unsupervised k-means algorithm. The codebook is learnt on a training set. The classification of a query image follows similar steps (Fig. 5). After extracting local image features and descriptors, the histogram representation is built up by using the codebook learnt on the training stage (Fig. 6). Our category model is simple: it is the sum of all histogram responses for each class (integrated response). As our category model yields acceptable results (Sec. 4), we refrain from using a classifier for building a model. The image window class is determined by finding the maxima of integrated responses of the three classes. For example, for the histogram representation shown Fig. 6a, the sum of all responses of Romanesque class is 5.6038, Gothic class – 1.8868 and Baroque class – 2.3019. Thus the image is classified as Romanesque. The histograms shown in Fig. 6 are built using a codebook of 30 cluster centers for each class. Note that for Romanesque class histogram high responses are located on the bins from 61 to 90, for Gothic class - from 1 to 30 and for Baroque class - from 31 to 60. The category model based on the maxima of the integraged class responses proves to be effective, as it makes the vote for the right class strong by integration of the high responses and suppresses the false class peaks, which may occur due to irrelevent descriptors located on architectural details, reflections and curtains.

4 Experiments of Window Classification and Discussion

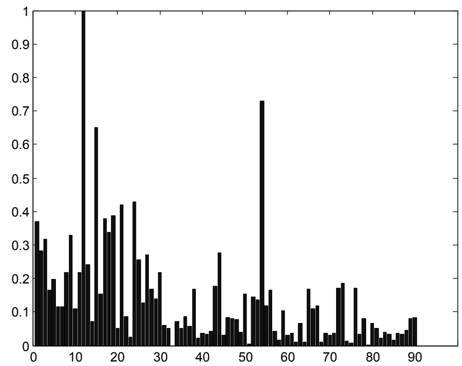
To the best knowledge of the authors there is no image database labeled by architectural styles. For testing and evaluation of our methodology we created a database of 400 images, 351 of which belong to our own and the rest - to Flickr[3] image datasets. 90 images of the database make the training set (1/3 of each class). The resolution range of the images is from 138×93 to 4320×3240 pixels.

To evaluate the issue of the codebook size (vocabulary size) we have perfomed an experiment with different codebook sizes (k) (Tab. 1 and Fig. 7). The value of peak threshold for SIFT feature extraction and the value of k for k-means clustering algorithm are searched so that the final classification rate is maximised on the training set. As it is obvious from Fig. 7, SIFT peak threshold values

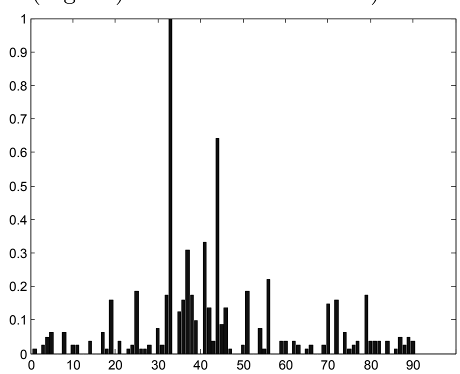
[3] http://www.flickr.com



a) Romanesque (Fig. 2a)



b) Gothic (Fig. 3a)



c)Baroque (Fig. 4c)

Fig. 6. Histograms of visual words for the images of different window styles.

larger than 0.03 decrease the classification rate. The reason for this is that the extraction of a bigger number of SIFT descriptors than that with peak threshold value equal to 0.03 tends to extract descriptors located on window reflections and backgound construction material textures, i.e. we are overfitting. Whereas peak threshold values smaller than 0.03 decrease the number of extracted SIFT

Table 1. Classification accuracy on the training set with different codebook sizes.

Peak Threshold (p)	k = 25	k = 30	k = 35	k = 40	k = 45
0,01	85,56	91,11	94,44	92,22	90,00
0,02	88,89	93,33	93,33	95,56	97,78
0,03	92,22	96,67	96,67	97,78	98,89
0,04	87,78	88,89	96,67	93,33	94,44
0,05	84,44	92,22	93,33	93,33	91,11

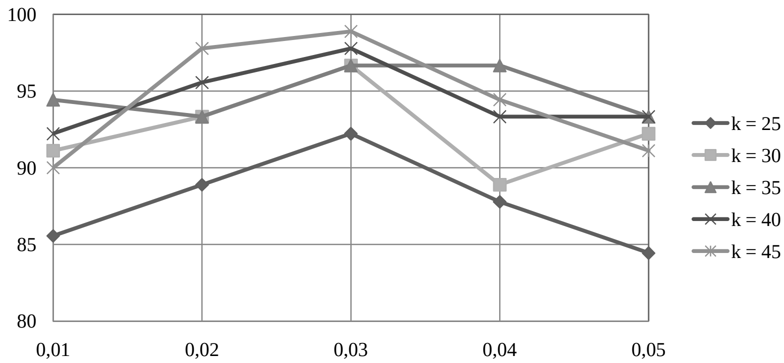


Fig. 7. Classification accuracy. Finding the best size of codebook (k) and SIFT peak threshold (p – horizontal axes).

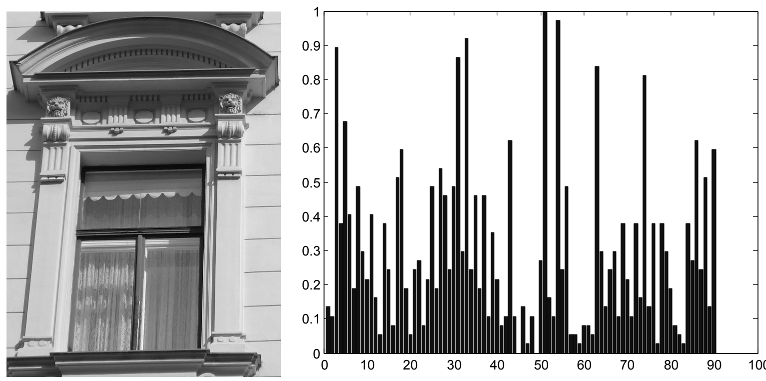
Table 2. Confusion matrix and the accuracy rate in parenthesis.

	Gothic	Baroque	Romanesque	Sum
Gothic	100 (98.1%)	1	1	102
Baroque	3	111 (92.5%)	6	120
Romanesque	1	3	84 (95.4%)	88
Sum	104	115	91	310

descriptors describing the dominating gradients characteristic for each window class. Fig. 7 also shows that the best choice for k-means algorithm k parameter is in the range $25-45$. We choose to take $k = 30$. The k parameter values smaller than 25 decrease the classification rate, as the number of cluster centers is not enough for the discrimination of visual words of different classes. Whereas values higher than 45 make the image histograms sparser, i.e. we get non-representative visual words. Our final codebook choice for testing the system is the one corresponding to $k = 30$ and peak threshold equal to 0.03. Running the classification with the mentioned codebook on a testing dataset of 310 images results in 15 false classified images, which yields an average classification rate of 95.16%. A confusion matrix, with true positives, is given in the Tab. 2.

In the Fig. 8 it is shown an example of a false classification of a Baroque window into Gothic. The sum of all responses of Romanesque class is 8.4324, Gothic class – 9.6757 and Baroque class – 8.7568. Therefore this image is classified as Gothic, since the maximum response is 9.6757. The reason for the false classification is the high complexity of architectural details and curtains.

As our approach uses SIFT features for classification, it is rotation and scale invariant [17]. The experiments also prove that the approach is camera viewpoint invariant, as the classification of windows is accurate under high perspective distortions.



a) Baroque window b) Histogram

Fig. 8. False classification of Baroque into Gothic window.

5 Conclusion

Virtual tourism, 3D building reconstruction, 3D city-modeling, CBIR and indexing of cultural heritage buildings operate on large image databases. The classification of such building databases into semantic categories belonging to certain architectural styles limits the image search of the whole databases to semantic portions. Also smart phones equipped with a building architectural style classification system may be applicable in the field of real tourism.

A method for window classification of Romanesque, Gothic and Renaissance/Baroque European main architectural styles was presented. In the scope of facade architectural style classification task by a voting mechanism of structural elements, like windows, domes, towers, columns, etc., the current paper purpose was to classify the architectural style taking into account only windows. Our approach is based on clustering and learning of local features. The experiments prove that the proposed approach yields a high classification rate.

Future work in the context of architectural style classification of building facades includes analysis of the images, which had a false classification due to high complexity of architectural details and curtains in order to eliminate false classifications, classification of windows on complete facade images, classification of other building structural elements, raising the number of classified architectural styles, use of symmetry feature descriptors and realization of a voting mechanism of different structural elements. The proposed methodology can be used for architectural style classification of other structural parts, like domes, towers, columns, etc.

References

1. Zheng, Y.T., Zhao, M., Song, Y., Adam, H., Buddemeier, U., Bissacco, A., Brucher, F., Chua, T.S., Neven, H.: Tour the world: building a web-scale landmark recogni-

tion engine. In: Proceedings of International Conference on Computer Vision and Pattern Recognition. (2009) 1085–1092
2. Zhang, W., Kosecka, J.: Hierarchical building recognition. Image and Vision Computing **25(5)** (2004) 704–716
3. Li, Y., Crandall, D., Huttenlocher, D.: Landmark classification in large-scale image collections. In: Proceedings of IEEE 12th International Conference on Computer Vision. (2009) 1957–1964
4. Cornelis, N., Leibe, B., Cornelis, K., Gool, L.V.: 3d urban scene modeling integrating recognition and reconstruction. International Journal of Computer Vision **78** (2008) 121–141
5. Snavely, N., Seitz, S.M., Szeliski, R.: Photo tourism: exploring photo collections in 3d. ACM Transaction on Graphics **25** (2006) 835–846
6. Ali, H., Seifert, C., Jindal, N., Paletta, L., Paar, G.: Window detection in facades. In: 14th International Conference on Image Analysis and Processing (ICIAP 2007), Springer-Verlag (2007) 837–842
7. Recky, M., Leberl, F.: Windows detection using k-means in cie-lab color space. In: International Conference on Pattern Recognition (ICPR 2010). (2010) 356–360
8. Recky, M., Leberl, F.: Window detection in complex facades. In: European Workshop on Visual Information Processing (EUVIP 2010). (2010) 220 – 225
9. Collins, P.: Changing Ideals in Modern Architecture, 1750-1950. McGill-Queen's University Press (1998)
10. Ojala, T., Pietikinen, M., Mäenpää., T.: Multiresolution grayscale and rotation invariant texture classification with local binary patterns. IEEE Trans. on Pattern Analysis and Machine Intelligence **24** (2002) 971–987
11. Haralick, R.M.: Statistical and structural approaches to texture. Proc. IEEE **67** (1979) 786–804
12. Zhang, D., Lu, G.: Review of shape representation and description techniques. Pattern Recognition **37** (2004) 1–19
13. Belongie, S., Malik, J., , Puzicha, J.: Shape matching and object recognition using shape contexts. IEEE Trans. on Pattern Analysis and Machine Intelligence **24** (2002) 509–522
14. Crowley, J.L., Parker, A.C.: A representation for shape based on peaks and ridges in the difference of lowpass transform. IEEE Trans. on Pattern Analysis and Machine Intelligence **6(2)** (1984) 156–170
15. Harris, C., Stephens, M.: A combined corner and edge detector. In: Proceedings of The Fourth Alvey Vision Conference. (1998) 147–151
16. Mikolajczyk, K., Schmid, C.: Indexing based on scale invariant interest points. In: Internationl Conference in Computer Vision. (2001) 525–531
17. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision **60(2)** (2004) 91–110
18. Matas, J., Chum, O., Urban, M., Pajdla1, T.: Robust wide baseline stereo from maximally stable extremal regions. In: British Machine Vision Conference. (2002) 384–393
19. Tuytelaars, T., Gool, L.V.: Wide baseline stereo matching based on local, affinely invariant regions. In: British Machine Vision Conference. (2000) 412–425
20. Csurka, G., Dance, C.R., Fan, L., Willamowski, J., Bray, C.: Visual categorization with bags of keypoints. In: Workshop on Statistical Learning in Computer Vision, ECCV. (2004) 1–22