

## 1. Introduction

In recent years there is a growing interest in the field of modeling, recognition and interpretation of urban environments. The main motivation for this study is to provide a tool which would help to annotate the windows (and later also other parts) on the facade images. Several types of users can benefit from such a tool: architects, analysts of architectural history, a visual navigation system, etc. Currently, for example, an architect has to manually annotate the windows, which is a tiresome and time consuming task for hundreds of facade images for a certain region of a larger city. The idea is to provide a semi-automatic tool which gives the first approximation of window locations, and users would only have to make the necessary corrections in case of a demand for greater accuracy. This would significantly reduce the annotation time required.

In this paper we focus on window recognition as a basic step in the description of facades. The recognition process starts from local structures at the pixel level. Most windows are rectangular in shape. Therefore, the linear edge segments in the facade images are a reasonable starting point for generating hypotheses about the position of windows on the facade. One can also use the fact that buildings are often divided into floors which are further divided into rooms. The room configuration is usually uniform for the whole building and, as a consequence, windows align horizontally (for individual floors) and vertically (for individual rooms on different floors). This is a key property of window candidates. Although they generally vary in shape and size, the windows on the same facade are often similar to each other, which can be used for recognition. In other words, multiple repetition of a certain part of the facade image increases the probability that this part corresponds to the window. Finally, it is possible to use machine learning methods with an appropriate set of features to develop a window detector based on the learned window appearance. Once the likely location of windows are known, one can generate a structural description of a facade that is interpretable and allows a further high-level processing.

The rest of the paper is structured as follows. In the next section we briefly present the related work. Section 3 describes the segmentation of input images which generates the hierarchical structure of window candidates which is then used in facade modeling via formal grammar rules. Section 4 provides the experimental results which are discussed in Section 5. The paper concludes in Section 6 where we give some ideas for further work.