An Edgel-Based Two Step Approach for Simultaneous Matching of Image Sequences

Andreas Klaus, Joachim Bauer and Konrad Karner
VRVis Research Center for Virtual Reality and Visualization - Institute for Computer Graphics and Vision, Graz University of Technology
{klaus,bauer,karner}@vrvis.at, http://www.vrvis.at/ar2/index.html

Abstract:
In this paper we report on a new edgel-based matching approach. High performance is reached using the gradient information of textured regions only and by applying a two step approach. The hierarchical iterative procedure provides fast convergence and avoids local minima. The method performs well in terms of accuracy, robustness and speed. Furthermore it allows to match an arbitrary number of oriented images simultaneously.

1 Introduction

Matching of image pairs and image sequences is a very important task in the computer vision community. Therefore a lot of effort has already been invested towards increasing accuracy, robustness and performance. The reasons why there is still a lot of research to do in this area, are on the one hand different requirements and on the other hand different intentions. Therefore we developed a completely new approach appropriate for our purpose, the reconstruction of facades from calibrated and oriented image sequences. A typical sequence is shown in Figure 1. Our requirements are a dense and accurate reconstruction as well as high performance. The latter is essential for our goal to enable the reconstruction of whole city districts from high resolution images within only a few days.

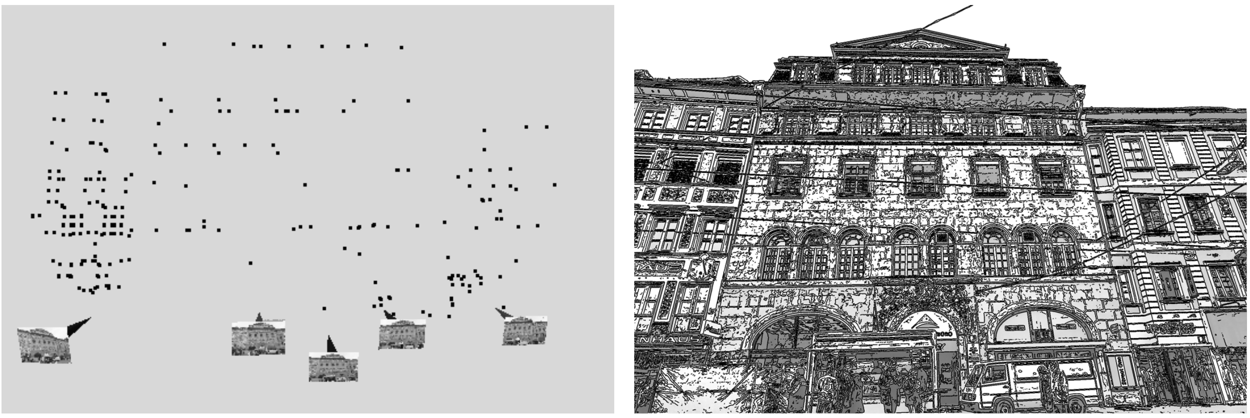
2 Our Work Flow

In this paper we are assuming that the relative or the geo-referenced orientation of the image sequence is already known. In order to determine the orientation very accurate, we are exploiting vanishing points and points of interest from line intersection [3]. Additional we utilize the homologous points of interest from the orientation process to derive a coarse depth map. This is achieved by performing a Delaunay triangulation from the resulting 3D points

(a) (b) (c) (d) (e)

Figure 1: A typical image sequence acquired at the Grazer Hauptplatz with a 3.3 megapixel digital consumer camera.

of interest. Figure 2a shows the reconstructed points of interest and the orientation of the cameras of the sequence shown in Figure 1. Our dense matching approach is performed in two consecutive steps, the first step preprocesses the data and transforms it in a suitable data structure. This allows a very efficient reconstruction step which is explained in Section 2.2.



(a) (b)

Figure 2: a) Orientation of the input image sequence shown in Figure 1 enhanced with the 3D-points of interest. b) Key image superimposed by the approximately 300000 extracted edgels.

2.1 Edgel extraction and creation of data structure

An important distinction to most aerial based matchers is that we are not using the intensity values directly. Instead of that we are focusing on the gradient information of textured regions. Therefore we are at first extracting edgels (edge elements) with the well known Canny method [1]. As a result we obtain the position with sub-pixel accuracy, the magnitude and the orientation of each edgel. In Figure 2b the extracted edgels of the image shown in Figure 1c are indicated as black points. Next we are exploiting the known epipolar geometry as proposed by Jung et al. [2] to determine candidates of homologous edgels. The main difference between both methods is that Jung uses chains to calculate 3D-point candidates whereas we

are using the edgels directly. Our candidates are calculated by projecting the edgels on the epipolar line.
We define one inner image of the sequence as key image. For each edgel in this image the line of sight is calculated and projected into the other images. This line of sight is limited in depth using the coarse 3D model from the orientation process and a fixed depth range. Only edgels that are very close to the corresponding epipolar line and with approximately same orientation as the edgel from the key image are considered as candidates. For each line of sight all candidates of each non-key image are recorded in a separate list, where we store the distance between the corresponding 3D-point and the origin of the key image in ascending order.

2.2 The surface reconstruction step

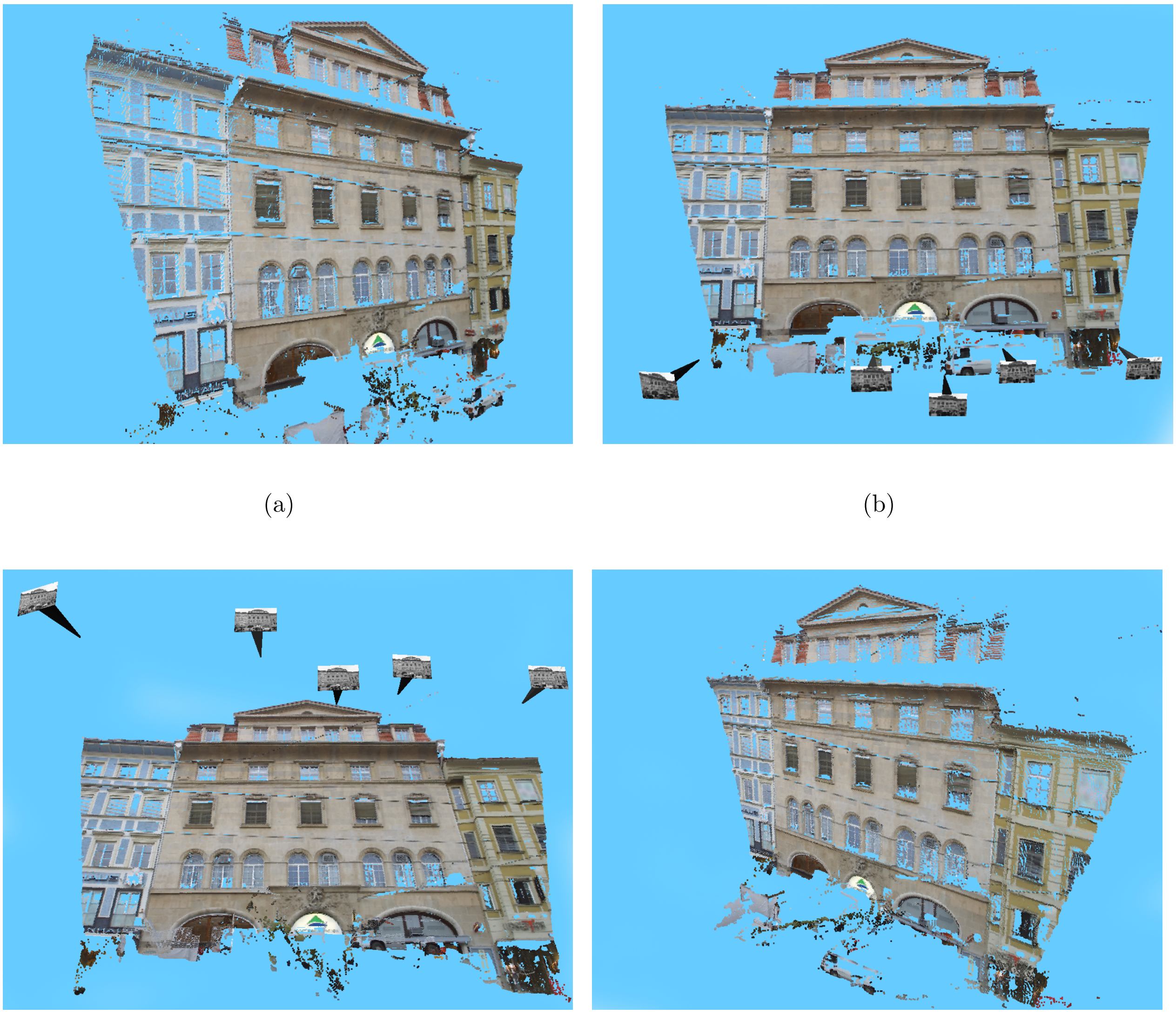
The task of this step is to find a rather smooth surface with as many 3D point candidates as possible in the close neighborhood. In our approach we focus on an iterative and hierarchical method to solve the corresponding problem inspired by a work published by Redert et al. [4]. The key image is subdivided into rectangles where the corner points form a regular grid. The calculation starts with a large grid size where the grid points are initialized by the coarse depth map. A refinement of the grid points is achieved iteratively by minimizing a cost function. For the determination of the cost function we are using the data structure we build in the previous step. It is therefore a very fast and simple task. For each edgel the actual depth is calculated by a bilinear interpolation between its grid points. The nearest candidates are searched for by exploiting the sorted lists and the accumulated distance is used to calculate the actual cost. The iterative refinement terminates if no further cost reduction is obtained. The grid size is halved and new grid points are initialized to proceed the refinement in the next finer level. This procedure is repeated until a required point density is reached. This hierarchical method convergences fast and avoids local minima especially when having repetitive structures within the images.

3 Performance and Experimental Results

task	result or additional information	execution time
Edgel extraction	ca. 275000 egels per image	16 sec
Creation of data structure	ca. 1.1 million sorted lists	29 sec
Surface reconstruction	ca. 2.4 million refinement steps	19 sec
all together	ca. 52000 calculated 3D-points	64 sec

Table 1: Performance for the image sequence of Figure 1.

Table 1 shows the performance of our approach tested on a Athlon PC with 1700 MHz. The task was to match the image sequence shown in Figure 1. The resolution of the five input images is 2160x1440 pixels. The most complex task is the creation of the data structure. Here we are using an optimized 3 dimensional kd-tree to reach high search performance. As result we obtain approximately 52000 3D-points that are shown in Figure 3.



(a) (b) (c) (d)

Figure 3: Reconstructed 3D points of the sequence in Figure 1 shown from different views.

4 Conclusions and Future Work

We have presented a new edgel-based matching approach that is particularly suitable for reconstruction of facades from calibrated and oriented image sequences. Therefore automatic reconstruction of man-made structures can be achieved using image sequences. This offers an alternative to the recent trend of laser scanning.

We are planning to extend our method to serve in other areas where fast, dense and accurate reconstruction is aspired. So far we obtain a 2 1/2 D reconstruction only. A full 3 D reconstruction might be achieved by repeating the mentioned method with other key images followed by a data linking step.

Acknowledgments

This work has been done in the VRVis research center, Graz and Vienna/Austria (http://www.vrvis.at), which is partly funded by the Austrian government research program Kplus.

References

[1] J. Canny. A computational approach to edge detection. In *RCV87*, pages 184–203, 1987.

[2] Franck Jung, Vincent Tollu, and Nicolas Paparoditis. Extracting 3d edgels hypotheses from multiple calibrated images: A step towards the reconstruction of curved and straight object boundary lines. *ISPRS Journal of Photogrammetric Computer Vision*, pages Part B 100–104, September 2002.

[3] Andreas Klaus, Joachim Bauer, and Konrad Karner. Metropogis: A semi-automatic city documentation system. *ISPRS Journal of Photogrammetric Computer Vision*, pages Part A 187–192, September 2002.

[4] Andre Redert, Emile Hendriks, and Jan Biemond. Correspondence estimation in image pairs. *IEEE Signal Processing Magazine*, pages 29–46, May 1999.