METROPOGIS: A CITY INFORMATION SYSTEM

K. Karner, J. Bauer, A. Klaus

VRVis Research Center
Graz, Austria

K. Schindler

Institute for Computer Graphics and Vision
Graz University of Technology, Austria

ABSTRACT

In this paper we report on a new system to augment a 3D block model of a real city obtained from aerial photogrammetry or aerial laser scanning with geo-referenced terrestrial data of the facades. The terrestrial images are acquired by a hand-held digital consumer camera. The relative orientation of the photographs is calculated automatically and fitted towards the 3D block model with minimized human input using vanishing points. The extraction of 3D primitives on the facades is based on line matching over multiple oriented images. The introduced city information system delivers a fully 3D geographic information data set and is called *MetropoGIS*.

1. INTRODUCTION

Virtual city models attract attention due to their usability for various applications such as the simulation of wind, flooding, traffic, and radio transmission as well as city planning, surveying, virtual tourism, cultural heritage documentation, and emergency training. For all of these applications different requirements on the data set have to be fulfilled. Important requirements are a fully three dimensional representation of the model with a high geometric accuracy and resolution. Furthermore, high resolution texture of all objects for photo-realistic visualization and a topological description of building features like windows and doors is necessary. In our system we attach importance to all these above mentioned requirements.

2. RELATED WORK

A working system for calibrated, terrestrial image acquisition in urban areas is described by Teller [1]. The system works with uncalibrated images and provides calibrated, georeferenced images as output. The usage of a GPS-based position estimation allows a fully automatic processing. The sensor provides omni-directional images for better pose estimation. Detailed information about this approach can be found in Antone [2], [3] and Bosse [4]. Our approach is quite similar but differs in some important aspects. One is that we can handle occlusions in a much easier way by evaluating multiple adjacent images simultaneous as described in Bornik [5]. In addition our input sensor, a digital camera (still or video) can be used in stop and go as well as in continuous mode. A GPS-based positioning system is not necessary and of limited use in narrow streets.

The 3D modeling of the facades is inspired by work done by Zisserman et al. [6], [7], [8] where an emphasis is put on the automatic extraction of planes from architectural images.

3. OVERVIEW OF OUR WORK FLOW

In our approach we assume that a 3D block model with known eave lines exists. We concentrate on the refinement of the facades of buildings using image sequences captured with a digital consumer camera from arbitrary positions. Our work flow consists of seven consecutive steps which will be explained in the following subsections.

3.1. Line Extraction and Vanishing Point Detection

Line extraction starts with an edge detection and linking process and yields contour chains with sub-pixel accuracy. For all contour chains of sufficient size a RANSAC [9] based line detection method is applied. The vanishing point detection is based on the method proposed by Rother [10].

3.2. Advanced Line and Point of Interest (POI) Extraction

The known position of vanishing points in the image is used to extract more lines pointing to these vanishing points. The extraction is based on a sweep line approach. In a preprocessing step edgels are extracted using a lower threshold than in 3.1. The amount of edgels to be processed is reduced by removing edgels with an orientation differing too much from the orientation to the vanishing point. The sweep line starts at the vanishing point and goes through the image plane. All edgels within some perpendicular distance to the sweep line are considered as inliers. Line segments are constructed by least squares fitting each densely chained subset of inliers. Overlapping parallel segments are merged after the sweep. In a post processing step intersections of line pairs from different vanishing points are computed. These intersections serve as points of interest (POI) for the computation of the relative orientation. Figure 1 shows the result of the advanced line extraction process.

Fig. 1. Detected line segments by the advanced extraction approach.

3.3. Relative Orientation of Image Pairs

In this approach we do not need time consuming point to point correlation in image pairs to find corresponding points. Instead we use vanishing points and POIs from line intersections to estimate the relative orientation. POIs are classified into 8 categories depending on the gradient information of lines and the position of the intersection relative to the lines (left or right, upper or lower corner). In Figure 2 POIs of different category are indicated by different colors. These POIs are accurate and invariant to perspective transformations.

Corresponding POIs in different images are found by testing a qualified set of possible POI pairs. For each potential pair the support is measured and the pair with the highest support is selected. The rotation of the left camera which is known from vanishing points is used to determine a plane, on which the POIs of the left image are projected as shown in Figure 3. A correct point pair is found by sampling through all possible combinations. If one potential point pair is assigned, the position of the second camera can only be shifted on the ray that goes through the corresponding 3D point on the plane, the camera center and the corresponding point in the right image plane. In Figure 3 this ray is indicated by the black continuous line.

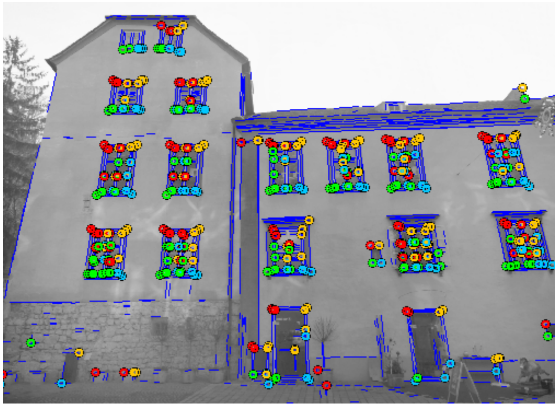


Fig. 2. Extracted POIs in one input image. The color of the detected POI indicates the category.

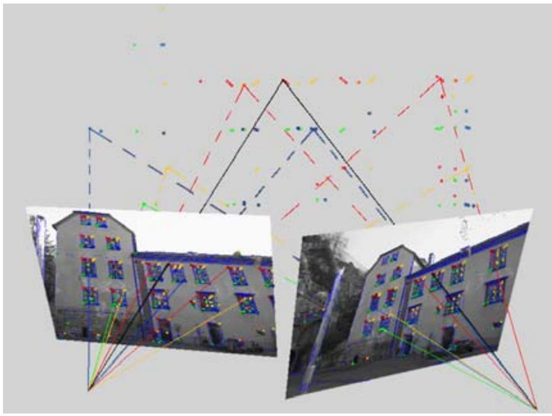


Fig. 3. Illustration how to determine relative orientation of an image pair using POIs.

If the right image is shifted along the corresponding ray the 3D points on the determined plane are projected onto different positions in the right image plane (see Figure 4). The final position is found by shifting the camera along the corresponding ray until the support is maximized.

3.4. Relative Orientation of Image Sequences

So far we have determined the orientation of image pairs and corresponding points in adjacent images. In order to calculate the orientation of a continuous sequence we start with one image pair and calculate 3D points assuming a fixed baseline (without loss of generality) and the rotation to the third image using the vanishing points. The position is found by selecting a 3D to 2D POI correspondence and shifting the third image along the obtained ray to maximize

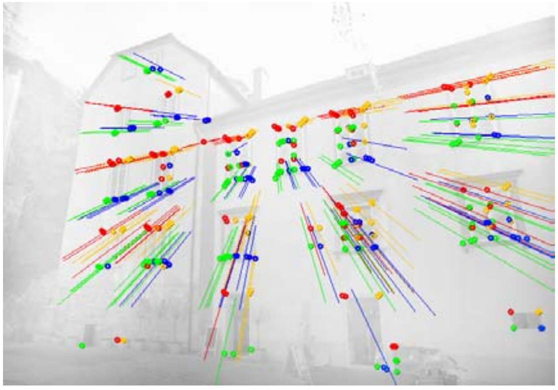


Fig. 4. If one corresponding point is given the second camera can only be shifted on the given ray.

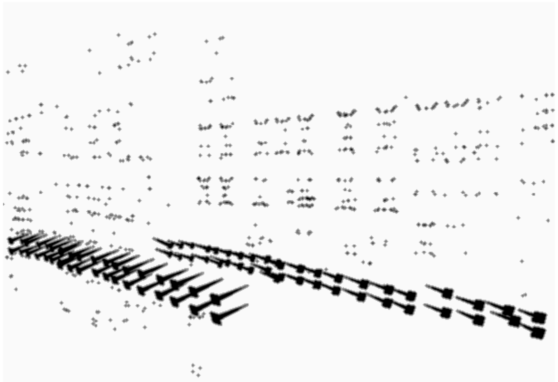


Fig. 5. Geo-referenced orientations of two image sequences of the Schlossbergplatz. Each camera is represented by a small plane and a direction arrow. At each location we shot two images with a short vertical baseline. The reconstructed POIs are indicated by dots.

the support of reprojected 3D points and 2D POIs in the third image. The corresponding points of the new image are used to calculate new and to improve old 3D points. This approach is repeated until all images are oriented to each other.

3.5. Image to 3D Model Fitting

So far we have only obtained a relative oriented sequence, where the position and orientation in geo-referenced coordinates as well as the scale are not yet known. Due to the fact that the vertical direction of an image is known from vanishing points only two control points are necessary to transform the image into a geo-referenced coordinate system. These two control points are taken from the known eave lines. In our approach only two images of a continuous sequence have to be fitted in such a semi-automatic way which minimizes the manual work. Figure 5 shows two sequences of images which were oriented semi-automatically.

3.6. Bundle Block Adjustment

For each building block a bundle adjustment is carried out to optimally fit the images into the existing block model. The known control points in some images help to fix the transformation to the world coordinate frame and stabilise the block, particularly in long sections of translational camera movement. Since the vertical direction is known from the detected vanishing points in the terrestrial images, a constraint is added to keep the mean vertical direction of the cameras unchanged.

3.7. Line Matching

The set of line segments per image together with the known orientation of the image sequence are the input for line matching. Our approach closely follows the one described by Schmid and Zisserman [11]. The result of the line matching process is a set of 3D lines in object space.

Basically the algorithm works as follows: For a reference line segment in one image of the sequence potential line matches in the other images are found by taking all lines that lie between the epipolar lines induced by the endpoints of the reference line segment.

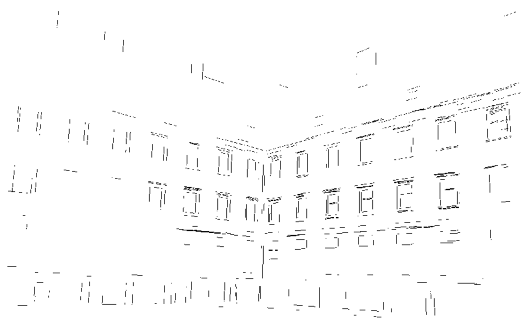
Each of these potentially corresponding line pairs gives a 3D line segment (except for those, which are parallel to the epipolar line, since in this case no intersection between the epipolar line and the image line can be computed).

The potential 3D lines are then projected into all remaining images. If image lines are found which are close to the reprojection, the candidate is confirmed, else it is discarded. Finally a correlation based similarity criterion is applied to select the correct line. Figure 6 shows two views of the extracted 3D line set. Obviously, due to the small vertical baseline the geometric accuracy of the horizontal line segments is limited.

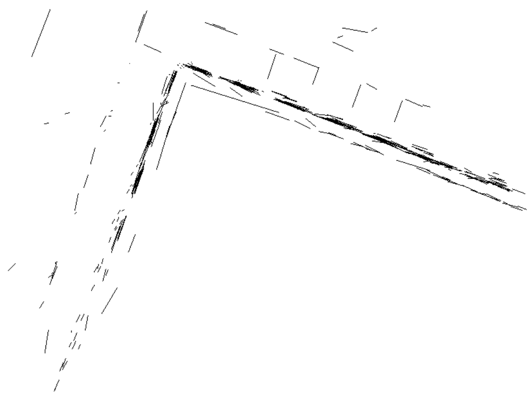
4. CONCLUSIONS AND FUTURE WORK

We have presented an image based information system called *MetropoGIS* in which a 3D block model of a city is augmented with terrestrial measured data. We have shown that it is possible to determine the geo-referenced orientation of a large image sequence with only a few mouse clicks without any position estimation. Because we use a consumer camera, our image acquisition is straightforward and allows high flexibility. By exploiting lines and vanishing points we have developed a robust and fast method to determinate relative orientation for even large baselines.

So far we have a semi-automatic system where an op-



(a)



(b)

Fig. 6. Two views of the extracted 3D line set of the facade in Figure 1. (a) front view (b) top view

erator is still involved in the work flow. We are planning to avoid this manual interaction by a fully automatic step where we try to match roof lines extracted from the terrestrial images with 3D lines from the block model. Furthermore, we are working on a plane sweeping approach to find object planes connected to 3D lines. Our approach is similar to the one presented in [6] with the difference that we use a feature based correlation criterion instead of area based cross correlation.

5. ACKNOWLEDGMENTS

Parts of this work have been done in the VRVis research center, Graz/Austria (http://www.vrvis.at), which is partly funded by the Austrian government research program Kplus. Other parts were supported by the EU IST project '3D Murale'.

6. REFERENCES

[1] S. Teller, "Calibrated, registered images of an extended urban area," in *IEEE Proceedings of CVPR*, 2001.

[2] M. Antone and S. Teller, "Automatic recovery of relative camera rotations for urban scenes.," in *CVPR*, 2000, pp. 282–289.

[3] M. Antone and S. Teller, "Scalable, absolute position recovery for omni-directional image networks.," in *CVPR*, 2001.

[4] M. Bosse, D. de Couto, and S. Teller, "Eyes of argus:georeferenced imagery in urban environments.," in *GPS World*, 1999, pp. 20–30.

[5] A. Bornik, K. Karner, J. Bauer, F. Leberl, and H. Mayer, "High-quality texture reconstruction from multiple views.," in *to appear in Journal of Visualization and Computer Animation*, 2002.

[6] A. Zisserman, T. Werner, and F. Schaffalitzky, "Towards automated reconstruction of architectural scenes from multiple images," in *Proc. 25th workshop of the Austrian Association for Pattern Recognition*, 2001, pp. 9–23.

[7] C. Schmid and A. Zisserman, "The geometry and matching of lines and curves over multiple views," *IJCV*, vol. 40, no. 3, pp. 199–233, December 2000.

[8] F. Schaffalitzky and A. Zisserman, "Planar grouping for automatic detection of vanishing lines and points," *IVC*, vol. 18, no. 9, pp. 647–658, June 2000.

[9] M. Fischler and R. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography.," *Communications of the Association for Computing Machinery, 24(6):381–395.*, 1981.

[10] C. Rother, "A new approach for vanishing point detection in architectural environments," in *Proceedings of the 11th British Machine Vision Conference*, 2000, pp. 382–391.

[11] C. Schmid and A. Zisserman, "Automatic line matching across views," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, 1997, pp. 666–671.