Generalized Use of Homographies for Piecewise Planar Reconstruction

Konrad Schindler

Computer Graphics and Vision
Graz University of Technology
Inffeldgasse 16, 8010 Graz, Austria
schindl@icg.tu-graz.ac.at

Abstract. We present a method for piecewise planar modeling from oriented images. The method uses a set of homologous image points and the underlying 3D geometry to determine image regions which satisfy plane-induced homographies. It robustly detects and reconstructs planes of arbitrary position and orientation in the scene and takes advantage of the regular raster structure of the images to delineate the regions.

1 Introduction

Automatic object reconstruction is a continuing goal of photogrammetry and computer vision. We address the task of reconstructing planar regions of a scene. Many man-made objects can be at least partly modeled with planes, and such a CAD-like representation is important for users of vision systems, as has been pointed out in [1]. Applications include surveying of buildings, documentation of cultural heritage [2] and reverse engineering tasks [3].

Techniques using inter-image homographies are particularly powerful, because they are based on the original gray-values and avoid influences of the point reconstruction process [4]. Previous methods using homographies are based on sweeping a plane along *one* degree of freedom, making it necessary to determine the other 3 degrees of freedom in advance. This is done with the help of reconstructed object lines, which are either used to determine the normal vectors of potential planes [4],[5] along which a translational plane sweep is performed, or a rotation axis for rotational sweeping [6]. We want to avoid this, firstly because it restricts the set of detectable planes to a certain parameter family and secondly because it relies on reconstructed object lines, which are often incomplete.

A problem of many reconstruction methods is to correctly delineate the planes. Using plane intersections, e.g. [4], is only possible, if one can be sure that the scene is composed entirely of planes and that all planes have been detected. Heuristic grouping rules similar to the ones used for delineation of roof polygons in aerial images, e.g. [7], are difficult to establish for more generic scenes.

In section 2 we will show that it is possible to use homographies to detect planes with arbitrary positions and orientations in space, if a dense set of corresponding image points can be established. We will also use 2D image-processing

J. Bigun and T. Gustavsson (Eds.): SCIA 2003, LNCS 2749, pp. 470–476, 2003.

techniques on the dense set of homologous points to find a coarse delineation of the planar regions. In section 3 we will present experimental results and in section 4 we will give a brief summary and outlook.

2 Planar Reconstruction

Previous sweeping methods establish the necessary correspondences between the images directly via the plane to be tested: gray-value edges are detected in one image with a point-of-interest operator and transformed to the other image with the homography induced by the plane [8]. The advantage of this method is that no feature correspondences have to be found in advance. A disadvantage is that correspondences are only found on edges, so that their density and distribution may vary from plane to plane. A second disadvantage is that at this stage it is not possible to decide, whether a point belongs to the potential planar region. This has to be overcome by a heuristic weighting procedure [8]. Thirdly the method is computationally expensive, because the correlation measures have to be recomputed for each homography.

2.1 Image Matching

To overcome the mentioned problems we use a dense matching algorithm to establish correspondences in advance. This guarantees uniform coverage of the whole image, since the locally best correspondence is found even in weakly textured regions. To be robust against totally untextured regions the matching algorithm uses a smoothness assumption, since most surfaces, which appear homogeneous under natural lighting conditions, are smooth [9]. For our modeling algorithm we need a regular raster of correspondences. Image matching is thus performed subject to the condition that homologous points lie on a regular raster in one of the images, which we will call the *master image* I_0. This can be achieved by fixing the raster points $\mathbf{x}_{i,0}$ in the master image and searching for the best correspondences $\mathbf{x}_{i,j}$ in the other n images.

Note that at the time of dense matching the epipolar geometry is already known, so that matching will in principle work even for wide baselines. However with increasing baseline self-occlusions lead to a growing number of false matches.

2.2 Homographies and Region Growing

The idea of using homographies for plane detection was first presented in [8]. A plane in 3D space, which is recorded with two pinhole cameras, defines a homography (a projective mapping) between the two recorded images: $\mathbf{x}_1 = \mathtt{H}\mathbf{x}_0$. For example, if the projection matrices of the two cameras are given by $\mathtt{C}_0 = [\mathtt{I}|\mathbf{0}]$ and $\mathtt{C}_1 = [\mathtt{A}|\mathbf{a}]$, respectively, and the plane is given by its homogeneous representation $\mathtt{P}$, then [10] the homography is given by[1]

[1] An expression for the homography between two arbitrary cameras is derived in [5].

$$\mathtt{H} = \mathtt{A} + \mathbf{a}\mathbf{v}^{\mathsf{T}} \quad \text{where} \quad \mathbf{v} = \frac{1}{p_4}(p_1, p_2, p_3)^{\mathsf{T}} \tag{1}$$

If we have recovered a metric reconstruction of the camera motion, we can use homographies to establish a RANSAC procedure for plane detection in the set of homologous points: in a metric framework three non-collinear object points uniquely define a plane, thus three non-collinear homologous points define a homography $\mathtt{H}_{(123)}$ between two images I_0 and I_1. Following the RANSAC-principle three such points forming an equilateral triangle are randomly selected and the homography $\mathtt{H}_{(123)}$ is computed. If the triangle is incident to a planar region of the object, all other homologous points $\mathbf{x}_i$ inside the triangle are subject to the same homography and must (up to a threshold, which depends on the matching accuracy) satisfy $\mathbf{x}_{i,1} = \mathtt{H}_{(123)}\mathbf{x}_{i,0}$. If this is the case, a planar triangle has been detected.

Since the homologous points lie on a regular grid in the master image, we can complete the region with a conventional 2D region-growing algorithm, using the triangle as seed region and the homography as homogeneity criterion. The described process is iterated, proceeding from bigger to smaller triangles until no more planes are found.

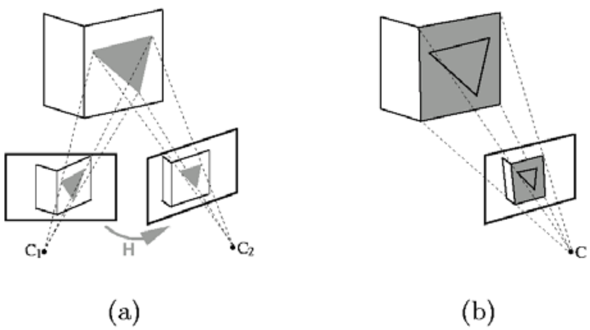
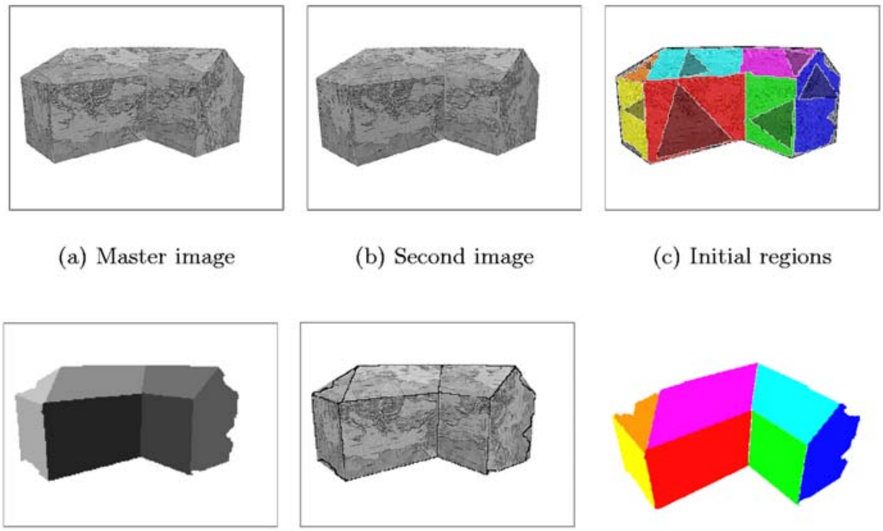


Fig. 1. Detection of planar regions. (a) The images of a planar triangle are related by a homography $\mathtt{H}$. (b) The triangle can be grown to find the image region which satisfies $\mathtt{H}$.

The method is related to region-based range image segmentation [11]. However it has the advantage that one does not need to consider the influences and correlations of the (overdetermined) depth reconstruction. In the case of more than two images the homography between the master image and each other image can be computed independently.

2.3 Plane Fitting

The detected planar regions are refined in a second pass. The homography for each region is recomputed with a robust estimation algorithm based on *all* its points. In principle any robust optimization algorithm could be used. Our implementation uses the least-median-of-squares (LMS) estimator [12], which has the advantage that one need not specify a threshold for separating inliers from outliers. With the enhanced homographies region growing is repeated starting from the triangles' centroids. Since at this stage all planes have already been detected, the borders between regions can be properly treated: points, which fulfill both homographies, are assigned to the region, where their residual is smaller.



(a) Master image (b) Second image (c) Initial regions

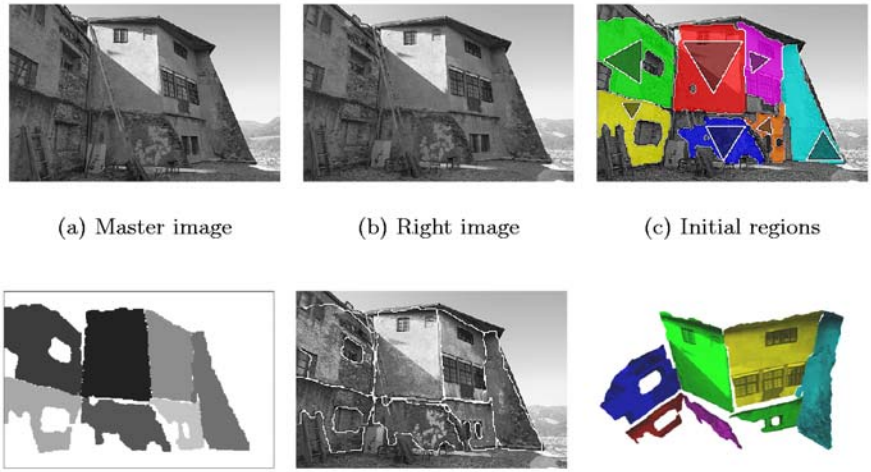
(d) Final regions (e) Region boundaries (f) 3D model

Fig. 2. Results for synthetic 'house' dataset. The errors near the silhouette are due to problems of the used image-matching algorithm at the borders between textured and untextured regions.

2.4 Delineation

Region growing –as opposed to testing all homologous points of the image– not only reduces computation time, but also guarantees that the regions lie within a closed delineation. Still one has to correct the boundaries between intersecting regions: The corresponding intersection lines are reprojected to the master image, where they are used to enhance the segmentation[2]. Ideally the rough delineation delivered by the region-growing scheme should be regarded as an additional source of information to be fused with heuristic delineation methods, which have been investigated in the context of aerial images [6],[7].

[2] Note that the complexity of the problem is lower than without previous segmentation, because the adjacency information from the 2D segmentation reduces the number of intersections to be tested.



(a) Master image (b) Right image (c) Initial regions

(d) Final regions (e) Region boundaries (f) 3D model

Fig. 3. Results for the ‘Oberkapfenberg castle’ dataset. At the time the images were taken reconstruction works were carried out, therefore the two lower left walls are partially occluded by tools and debris.

3 Experimental Results

To illustrate the functioning of the described algorithm, experiments on synthetic as well as real data are presented.

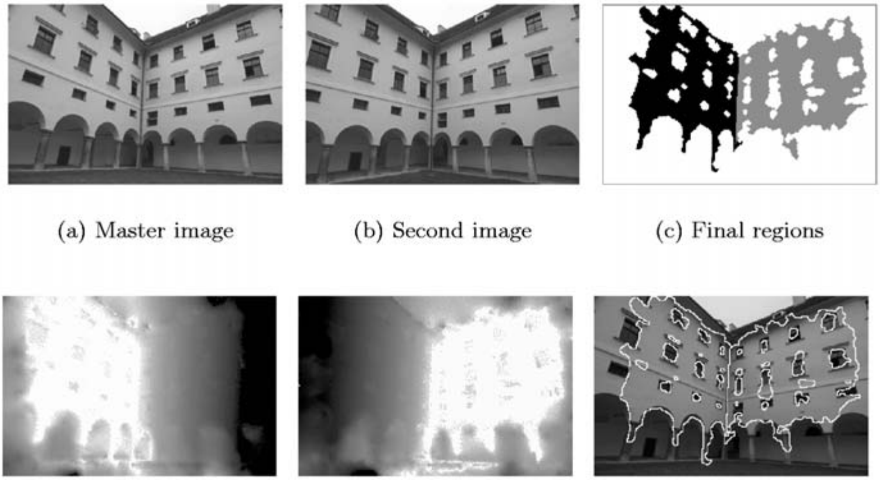
3.1 Synthetic Images

The dataset consists of two images with resolution 1000×750 pixels rendered from a simple CAD-model of a house, which was textured with a map of the Mediterranean. A regular grid in the master image with a grid spacing of 4 pixels was matched and the algorithm described in the previous sections was used to segment the dataset. The results are summarized in Figure 2.

3.2 Real Images

For the first example an image triple of the corner tower of the ’Oberkapfenberg castle’ near Graz was recorded with a calibrated Canon D30 (resolution 2160×1440 pixels) and oriented. The central image served as master image. A regular grid in the master image with a grid spacing of 8 pixels was matched. The segmentation results are shown in Figure 3.

The second example is a corner in the inner court of the ’Minorite Monastery’ in Graz, recorded with the same camera and oriented. Again a regular grid with a



(a) Master image (b) Second image (c) Final regions

(d) 1st homography (e) 2nd homography (f) Region boundaries

Fig. 4. Results for the 'Minorite Monastery' dataset. The two lower left images show gray-value plots of the residuals w.r.t. the two detected homographies (the darker a pixel is, the higher is its residual). On the right wall one can see the limitations of the smoothness assumption: in some regions without texture the matching is not correct, leading to holes in the recovered wall plane.

grid spacing of 8 pixels was matched. The segmentation results are shown in Figure 4. The example also illustrates how large untextured regions can deteriorate the result.

4 Concluding Remarks

We have presented a new method for the reconstruction of object planes from images, which uses plane-induced homographies to recover planar regions of arbitrary position and orientation. The method is a generalization of previous sweeping methods based on homographies, which search the position of a plane along one degree of freedom. In order to be efficient, a dense matching between images is established in advance and corresponding points are used to verify or discard potential planes via the homographies they define between the images. Region growing is used to find a rough delineation of each detected region. We see the presented method as an additional way of planar reconstruction, not as a substitution for existing algorithms. An integration of different methods would be desirable to build a robust system.

An improvement could be to use the detected planes and associated homographies to upgrade the homologous points and and recover additional 3D points. The main problem of the current implementation is that the resulting model of the scene is effectively $2\frac{1}{2}$-dimensional, because it is based on only one master image. We believe that it will be possible to merge several partial $2\frac{1}{2}$D models based on the rich literature about modeling from multiple range images, e.g. [13].

Acknowledgments. This work has been supported by the European Commission under contract IST-1999-20273. The author would like to thank Markus Grabner for the 'virtual turntable' software, Konrad Karner for the 'Oberkapfenberg' pictures and Andreas Klaus for orienting the 'Minorite Monastery' sequence.

References

1. van den Heuvel, F.: Trends in CAD-based photogrammetric measurement. International Archives of Photogrammetry and Remote Sensing **33** (2000) 852–863
2. Wiedemann, A.: Digital architectural photogrammetry for building registration. In: Proc. International Kolloquium on Applications of Computer Science and Mathematics in Architecture and Civil Engineering (IKM97), Weimar. (1997)
3. Petran, F., Krzystek, P., Bonitz, P.: CAD-based reverse engineering with digital photogrammetry. In: International Archives Of Photogrammetry and Remote Sensing. Volume 31/5. (1996) 475–480
4. Zisserman, A., Werner, T., Schaffalitzky, F.: Towards automated reconstruction of architectural scenes from multiple images. In: Proc. 25th workshop of the Austrian Association for Pattern Recognition, Berchtesgaden, Germany. (2001) 9–23
5. Bauer, J., Karner, K., Schindler, K.: Plane parameter estimation by edge set matching. In: Proc. 26th workshop of the Austrian Association for Pattern Recognition, Graz, Austria. (2002) 29–36
6. Baillard, C., Zisserman, A.: A plane-sweep strategy for the 3d reconstruction of buildings from multiple images. In: International Archives of Photogrammetry and Remote Sensing. Volume 32/2. (2000) 56–62
7. Noronha, S., Nevatia, R.: Detection and description of buildings from multiple aerial images. In: Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR'97). (1997) 588–594
8. Baillard, C., Schmid, C., Zisserman, A., Fitzgibbon, A.: Automatic line matching and 3d reconstruction of buildings from multiple views. In: Proc. ISPRS Conference on Automatic Extraction of GIS Objects from Digital Imagery, Munich. (1999) 69–80
9. Kraus, K.: Photogrammetrie, Band 1. Dümmler Verlag (1994)
10. Luong, Q., Vieville, T.: Canonical representations for the geometries of multiple projective views. Computer Vision and Image Understanding **64** (1996) 193–229
11. Besl, P., Jain, R.: Range image segmentation. In: Machine Vision: Algorithms, Architectures, and Systems. Academic Press Professional, Inc. (1988) 221–256
12. Rousseeuw, P., Leroy, A.: Robust Regression and Outlier Detection. John Wiley and Sons (1987)
13. Hoover, A., Goldgof, D., Bowyer, K.: The space envelope: A representation for 3d scenes. Computer Vision and Image Understanding **69** (1998) 310–329